GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:34 ; Search time 3675.59 Seconds
(without alignments)
8198.470 Million cell updates/sec

Title: US-09-497-957-9
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1440	100.0	1440	6	AR117793	AR117793 Sequence
2	1440	100.0	1440	6	AR149463	AR149463 Sequence
3	1440	100.0	2727	9	HSU60319	U60319 Homo sapien
4	1438.4	99.9	1440	6	AR117794	AR117794 Sequence
5	1438.4	99.9	1440	6	AR117795	AR117795 Sequence
6	1438.4	99.9	1440	6	AR149464	AR149464 Sequence
7	1438.4	99.9	1440	6	AR149465	AR149465 Sequence
8	1436.8	99.8	1440	6	AR117796	AR117796 Sequence
9	1436.8	99.8	1440	6	AR149466	AR149466 Sequence
10	1201	83.4	1280	9	HSA249335	AJ249335 Homo sapi
11	1189	82.6	1200	9	AF115265	AF115265 Homo sapi
12	993	69.0	1045	9	AF079407	AF079407 Homo sapi
13	897	62.3	1320	4	AY007541	AY007541 Ceratothe
14	897	62.3	1320	4	AY007543	AY007543 Dicerorhi
15	896.8	62.3	1319	4	AY007544	AY007544 Rhinocero
16	893.8	62.1	2332	4	AY007542	AY007542 Diceros b
17	849.4	59.0	1085	9	HSA249336	AJ249336 Homo sapi
18	787	54.7	1073	9	HSA249337	AJ249337 Homo sapi
19	775.2	53.8	1885	9	AF144242	AF144242 Homo sapi
20	711.4	49.4	823	9	AF079408	AF079408 Homo sapi
21	637.6	44.3	1706	10	RNHERHAEM	AJ001517 Rattus no
22	622.8	43.2	1529	10	MMU66849	U66849 Mus musculu
23	617.4	42.9	781	9	AF079409	AF079409 Homo sapi
24	573	39.8	809	9	HSA250635	AJ250635 Homo sapi
25	549.8	38.2	987	9	AF150664	AF150664 Homo sapi
26	486.8	33.8	686	9	AF144240	AF144240 Homo sapi
27	478.2	33.2	789	10	AF008587	AF008587 Rattus no
28	464.8	32.3	804	9	AF149804	AF149804 Homo sapi
29	448.8	31.2	819	10	AF176534	AF176534 Rattus no
30	436.4	30.3	735	9	AF115264	AF115264 Homo sapi
31	393	27.3	560	9	AF144243	AF144243 Homo sapi
32	322	22.4	874	9	HSHLAH2	Y09800 H.sapiens H
33	321	22.3	10825	6	AR117789	AR117789 Sequence
34	321	22.3	10825	6	AR117790	AR117790 Sequence
35	321	22.3	10825	6	AR149459	AR149459 Sequence
36	321	22.3	10825	6	AR149460	AR149460 Sequence
37	321	22.3	12146	9	HSHFE	Z92910 Homo sapien
c 38	321	22.3	193752	2	AL359892	AL359892 Homo sapi
39	321	22.3	246240	6	AR036572	AR036572 Sequence
40	321	22.3	246240	6	AR036573	AR036573 Sequence
41	321	22.3	246240	6	AR036574	AR036574 Sequence
c 42	321	22.3	246282	9	HSU91328	U91328 Human hered
43	320.4	22.2	492	9	AF144241	AF144241 Homo sapi
44	319.4	22.2	10825	6	AR117791	AR117791 Sequence
45	319.4	22.2	10825	6	AR117792	AR117792 Sequence

ALIGNMENTS

RESULT 1
AR117793
LOCUS AR117793 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 9 from patent US 6140305.
ACCESSION AR117793
VERSION AR117793.1 GI:14098699
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 9 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. .1440
/organism="unknown"
BASE COUNT 347 a 355 c 407 g 331 t
ORIGIN

Query Match 100.0%; Score 1440; DB 6; Length 1440;

```
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   2
AR149463
LOCUS       AR149463                1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 9 from patent US 6228594.
ACCESSION   AR149463
VERSION     AR149463.1  GI:15114054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 9 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    355 c    407 g    331 t
ORIGIN


  Query Match             100.0%;  Score 1440;  DB 6;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
```

```
Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
```

```
Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT   3
HSU60319
LOCUS       HSU60319                2727 bp    mRNA    linear   PRI 29-OCT-1997
DEFINITION  Homo sapiens haemochromatosis protein (HLA-H) mRNA, complete cds.
ACCESSION   U60319
VERSION     U60319.1  GI:1469789
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2727)
  AUTHORS   Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A.,
            Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A.,
            Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P.,
            Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C.,
            Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E.,
            Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J.,
            Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
  TITLE     A novel MHC class I-like gene is mutated in patients with
            hereditary haemochromatosis
  JOURNAL   Nature Genet. 13 (4), 399-408 (1996)
  MEDLINE   96331279
REFERENCE   2  (bases 1 to 2727)
  AUTHORS   Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A.,
            Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A.,
            Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P.,
            Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C.,
            Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E.,
            Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J.,
            Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1996) Mercator Genetics, 4040 Campbell Ave.,
            Menlo Park, CA 94025, USA
FEATURES             Location/Qualifiers
     source          1. .2727
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.3"
     gene            1. .2727
                     /gene="HLA-H"
                     /note="HFE"
     CDS             222. .1268
                     /gene="HLA-H"
                     /codon_start=1
                     /product="haemochromatosis protein"
                     /protein_id="AAC51823.1"
                     /db_xref="GI:1469790"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                     SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                     EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                     SRGAMGHYVLAERE"
BASE COUNT      702 a    606 c    660 g    759 t
ORIGIN
```

```
  Query Match           100.0%;  Score 1440;  DB 9;  Length 2727;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
```

```
Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

RESULT 4
AR117794
LOCUS AR117794 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 10 from patent US 6140305.
ACCESSION AR117794
VERSION AR117794.1 GI:14098700
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 10 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. .1440
/organism="unknown"
BASE COUNT 348 a 355 c 406 g 331 t
ORIGIN

Query Match 99.9%; Score 1438.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
```

```
Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

RESULT 5
AR117795
LOCUS AR117795 1440 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 11 from patent US 6140305.
ACCESSION AR117795
VERSION AR117795.1 GI:14098701
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 11 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. .1440
/organism="unknown"
BASE COUNT 347 a 354 c 408 g 331 t
ORIGIN

Query Match 99.9%; Score 1438.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT   6
AR149464
LOCUS       AR149464                1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 10 from patent US 6228594.
ACCESSION   AR149464
VERSION     AR149464.1  GI:15114055
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 10 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    355 c    406 g    331 t
ORIGIN
```

```
  Query Match             99.9%;  Score 1438.4;  DB 6;  Length 1440;
  Best-Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
```

```
Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   7
AR149465
LOCUS       AR149465                1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 11 from patent US 6228594.
ACCESSION   AR149465
VERSION     AR149465.1  GI:15114056
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 11 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    354 c    408 g    331 t
ORIGIN


  Query Match             99.9%;  Score 1438.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   8
AR117796
LOCUS       AR117796               1440 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 12 from patent US 6140305.
ACCESSION   AR117796
VERSION     AR117796.1  GI:14098702
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 12 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    354 c    407 g    331 t
ORIGIN


  Query Match             99.8%;  Score 1436.8;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1438;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   9
AR149466
LOCUS       AR149466               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 12 from patent US 6228594.
ACCESSION   AR149466
VERSION     AR149466.1  GI:15114057
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 12 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    354 c    407 g    331 t
```

```
ORIGIN

  Query Match            99.8%;  Score 1436.8;  DB 6;  Length 1440;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1438;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT  10
HSA249335
LOCUS       HSA249335               1280 bp    mRNA    linear   PRI 04-SEP-2001
DEFINITION  Homo sapiens mRNA for hemochromatosis protein (HFE gene) splice
            variant 1.
ACCESSION   AJ249335
VERSION     AJ249335.1  GI:15485418
KEYWORDS    alternative splicing; hemochromatosis protein; HFE gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1280)
  AUTHORS   OIiva,R. and Sanchez,M.
  TITLE     Identification of different alternative splicing forms of the HFE
            gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1280)
  AUTHORS   Oliva,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-SEP-1999) Oliva R., Faculty of Medicine and Clinic
            Hospital, Human Genome Research Group, Casanova 143, 08036, SPAIN
FEATURES             Location/Qualifiers
     source          1. .1280
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22"
                     /cell_line="HepG2"
     gene            161. .1138
                     /gene="HFE"
     CDS             161. .1138
                     /gene="HFE"
                     /function="iron metabolism"
                     /note="alternative splicing form with deletion of first 69
                     pb of exon 2"
                     /codon_start=1
                     /evidence=experimental
                     /product="hemochromatosis protein"
                     /protein_id="CAC67792.1"
                     /db_xref="GI:15485419"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLPLGYVDDQLFVFYDHESRR
```

```
                     VEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCE
                     MQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYL
                     ERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLK
                     DKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWEP
                     SPSGTLVIGVISGIAVFVVILFIGILFIILRKRQGSRGAMGHYVLAERE"
BASE COUNT      311 a    314 c    371 g    284 t
ORIGIN


  Query Match           83.4%;  Score 1201;  DB 9;  Length 1280;
  Best Local Similarity  94.9%;  Pred. No. 0;
  Matches 1280;  Conservative    0;  Mismatches    0;  Indels   69;  Gaps    1;

Qy      62 CTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAGC 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAGC 60

Qy     122 CTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGA 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGA 120

Qy     182 AGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCC 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCC 180

Qy     242 GGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTC 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGC---- 236

Qy     302 ACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTT 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     237 ------------------------------------------------------------ 236

Qy     362 TGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCG 421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     237 -----CTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCG 291

Qy     422 TGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCT 481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 TGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCT 351

Qy     482 GAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGA 541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     352 GAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGA 411

Qy     542 AAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCA 601
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     412 AAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCA 471

Qy     602 AGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGA 661
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     472 AGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGA 531

Qy     662 ATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCT 721
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     532 ATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCT 591

Qy     722 GGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTG 781
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     592 GGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTG 651

Qy     782 CCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCC 841
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     652 CCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCC 711

Qy     842 TCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGC 901
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     712 TCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGC 771

Qy     902 CTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGA 961
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     772 CTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGA 831

Qy     962 TGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTG 1021
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     832 TGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTG 891

Qy    1022 GATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCC 1081
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     892 GATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCC 951

Qy    1082 AGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCAT 1141
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     952 AGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCAT 1011

Qy    1142 TGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCAT 1201
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1012 TGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCAT 1071

Qy    1202 AATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACG 1261
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1072 AATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACG 1131

Qy    1262 TGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA 1321
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1132 TGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA 1191

Qy    1322 GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATT 1381
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1192 GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATT 1251

Qy    1382 GCCTGACGAACTCCTTGATTTTAGCCTTC 1410
           |||||||||||||||||||||||||||||
Db    1252 GCCTGACGAACTCCTTGATTTTAGCCTTC 1280


RESULT  11
AF115265
LOCUS       AF115265                1200 bp    mRNA    linear   PRI 07-MAY-2001
DEFINITION  Homo sapiens hemochromatosis termination variant terE6 (HFE) mRNA,
            complete cds.
ACCESSION   AF115265
VERSION     AF115265.1  GI:11094314
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
  AUTHORS   Thenie,A., Orhant,M., Gicquel,I., Fergelot,P., Le Gall,J.Y.,
            David,V. and Mosser,J.
  TITLE     The HFE gene undergoes alternate splicing processes
  JOURNAL   Blood Cells Mol. Dis. 26 (2), 155-162 (2000)
  MEDLINE   20448010
   PUBMED   11001625
REFERENCE   2  (bases 1 to 1200)
  AUTHORS   Thenie,A., Orhant,M. and Mosser,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-1998) UPR 41 CNRS, Faculte de Medecine, 2, av du
            Pr. Bernard, Rennes 35043, France
FEATURES             Location/Qualifiers
     source          1. .1200
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22"
     gene            1. .1200
                     /gene="HFE"
     CDS             12. .1058
                     /gene="HFE"
                     /codon_start=1
                     /product="hemochromatosis termination variant terE6"
                     /protein_id="AAG29572.1"
                     /db_xref="GI:11094315"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
```

```
                 WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                 RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                 SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                 EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                 SRGAMGHYVLAERE"
BASE COUNT      298 a    290 c    346 g    266 t
ORIGIN


  Query Match           82.6%;  Score 1189;  DB 9;  Length 1200;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1189;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    211 GAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGA 270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGA 60

Qy    271 CCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTG 330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTG 120

Qy    331 CCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGC 390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGC 180

Qy    391 TGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCA 450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCA 240

Qy    451 GTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACA 510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACA 300

Qy    511 TGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACA 570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACA 360

Qy    571 CCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGA 630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGA 420

Qy    631 AGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAG 690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAG 480

Qy    691 CAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCA 750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCA 540

Qy    751 GGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGC 810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGC 600

Qy    811 TGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGA 870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGA 660

Qy    871 CCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCA 930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCA 720

Qy    931 TGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTAT 990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTAT 780

Qy    991 TGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAG 1050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAG 840

Qy   1051 AGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCT 1110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCT 900

Qy   1111 GGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTG 1170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTG 960

Qy   1171 TCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAG 1230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAG 1020

Qy   1231 GAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTG 1290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTG 1080

Qy   1291 TGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTT 1350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTT 1140

Qy   1351 CAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGA 1399
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGA 1189


RESULT  12
AF079407
LOCUS       AF079407                1045 bp    mRNA    linear   PRI 18-MAR-1999
DEFINITION  Homo sapiens hemochromatosis splice variant del14E4 (HFE) mRNA,
            complete cds.
ACCESSION   AF079407
VERSION     AF079407.1  GI:3695106
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1045)
  AUTHORS   Rhodes,D.A. and Trowsdale,J.
  TITLE     Alternate splice variants of the hemochromatosis gene Hfe
  JOURNAL   Immunogenetics 49 (4), 357-359 (1999)
  MEDLINE   99180629
REFERENCE   2  (bases 1 to 1045)
  AUTHORS   Rhodes,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-1998) Immunology, University of Cambridge, Tennis
            Court Road, Cambridge CB2 1QP, UK
FEATURES             Location/Qualifiers
     source          1. .1045
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1"
     gene            1. .1045
                     /gene="HFE"
     CDS             37. .1041
                     /gene="HFE"
                     /codon_start=1
                     /product="hemochromatosis splice variant del14E4"
                     /protein_id="AAC62646.1"
                     /db_xref="GI:3695107"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVTTLRCRALNYYP
                     QNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLD
                     QPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQGSRGAMGHYVLAERE
                     "
BASE COUNT      243 a    259 c    314 g    229 t
ORIGIN


  Query Match           69.0%;  Score 993;  DB 9;  Length 1045;
  Best Local Similarity  96.1%;  Pred. No. 2.7e-280;
  Matches 1045;  Conservative    0;  Mismatches    0;  Indels   42;  Gaps    1;

Qy    186 GAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCG 245
```

```
Db      1 GAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCG 60

Qy    246 CTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACAC 305
Db     61 CTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACAC 120

Qy    306 TCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAA 365
Db    121 TCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAA 180

Qy    366 GCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTG 425
Db    181 GCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTG 240

Qy    426 GAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGT 485
Db    241 GAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGT 300

Qy    486 CAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAAT 545
Db    301 CAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAAT 360

Qy    546 CACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAA 605
Db    361 CACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAA 420

Qy    606 GACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTC 665
Db    421 GACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTC 480

Qy    666 TGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAG 725
Db    481 TGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAG 540

Qy    726 TGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCT 785
Db    541 TGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCT 600

Qy    786 GCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCT 845
Db    601 GCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACA---------- 650

Qy    846 TTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTG 905
Db    651 ------------------------------AGTGACCACTCTACGGTGTCGGGCCTTG 678

Qy    906 AACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCC 965
Db    679 AACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCC 738

Qy    966 AAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATA 1025
Db    739 AAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATA 798

Qy   1026 ACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGC 1085
Db    799 ACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGC 858

Qy   1086 CTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGA 1145
Db    859 CTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGA 918

Qy   1146 GTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATA 1205
Db    919 GTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATA 978

Qy   1206 TTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAG 1265
Db    979 TTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAG 1038

Qy   1266 TGACACG 1272
Db   1039 TGACACG 1045
```

```
RESULT 13
AY007541
LOCUS       AY007541               1320 bp    mRNA    linear   MAM 22-OCT-2000
DEFINITION  Ceratotherium simum HFE protein mRNA, complete cds.
ACCESSION   AY007541
VERSION     AY007541.1  GI:10945687
KEYWORDS    .
SOURCE      white rhinoceros.
  ORGANISM  Ceratotherium simum
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
REFERENCE   1  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Ceratotherium simum"
                     /db_xref="taxon:9807"
     CDS             86. .1132
                     /codon_start=1
                     /product="HFE protein"
                     /protein_id="AAG23701.1"
                     /db_xref="GI:10945688"
                     /translation="MGPRARPALFFLILLRTVAAQGRPPRSHSLRYLFMGASERDHGL
                     PLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWLQLSQSLKGWDHMFIVDF
                     WTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDHLEFCPETLDWRAAES
                     RALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQVPPLVKVTHHVAS
                     AVTTLRCQALNFYPQNITMRWLKDRKPMDVKDAESKDVLPSGDGTYQSWEALAVPPGE
                     EQRYTCQVEHPGLDQPLTATWEPSLSNTLVTGVISGIAVCVIIFFIGILFRILRKRQA
                     SRGAMGDYVLGECE"
BASE COUNT      284 a    363 c    394 g    279 t
ORIGIN
```

```
  Query Match           62.3%;  Score 897;  DB 4;  Length 1320;
  Best Local Similarity  82.4%;  Pred. No. 4.5e-252;
  Matches 1065;  Conservative    0;  Mismatches  225;  Indels    3;  Gaps    3;

Qy    137 AGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACG 196
Db      2 AGAGTGACTTCTTGGGTCACTCCTATTTCCCCGCCCCACCGAAGCAGCTGA-CGCGGGCC 60

Qy    197 GGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCT 256
Db     61 AGCACCCGCGGTCCGAGCCGGGGAAATGGGCCCGCGAGCCCGACCCGCGCTCTTCTTCCT 120

Qy    257 GATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTA 316
Db    121 GATCCTCCTGCGGACCGTGGCCGCGCAGGGGCGACCACCGCGGTCACACTCTCTGCGCTA 180

Qy    317 CCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTA 376
Db    181 CCTCTTCATGGGCGCCTCAGAGCGGGACCACGGCCTGCCCCTGTTTGAAGCCTTGGGCTA 240

Qy    377 CGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAAC 436
Db    241 TGTGGATGACGAGCTTTTCGTGGCCTACAATCACGAGAGTCGCCGTGCAGAGTCCCGTGC 300

Qy    437 TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA 496
Db    301 CCAGTGGGTCTTGGGTGAGGCCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAA 360

Qy    497 AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAG 556
```

```
Db    361 AGGGTGGGATCACATGTTCATCGTCGACTTCTGGACCATCATGGACAACCACAACCACAG 420

Qy    557 CAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTAC 616

Db    421 CAAGGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGGTGCAAGAGGACAACAGTAC 480

Qy    617 CGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACAC 676

Db    481 CAGGGGTTTCTGGAAGTACGGGTACGATGGGCAGGACCATCTTGAATTCTGCCCTGAGAC 540

Qy    677 ACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCA 736

Db    541 ACTGGATTGGAGAGCAGCAGAGTCCAGGGCCCTGACCACCAAGCTGGAGTGGGAAGTGAA 600

Qy    737 CAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCA 796

Db    601 CAAGATTCGGGCCAAGCAGAACAGAGCCTATCTCGAGAGGGACTGCCCTGAGCAGCTGCA 660

Qy    797 GCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGT 856

Db    661 GTGGTTGCTGGAGCTGGGGAGAGGGGTTCTGGACCAGCAAGTGCCTCCTTTGGTGAAGGT 720

Qy    857 GACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCC 916

Db    721 GACTCATCATGTGGCCTCTGCAGTGACCACTCTACGGTGTCAGGCTCTGAACTTCTACCC 780

Qy    917 CCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGA 976

Db    781 CCAGAACATCACCATGAGGTGGTTGAAGGACAGGAAGCCAATGGATGTCAAGGACGCCGA 840

Qy    977 ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGT 1036

Db    841 GTCTAAGGATGTGCTGCCCAGCGGGGACGGGACCTACCAGAGCTGGGAGGCTTTGGCCGT 900

Qy   1037 ACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 1096

Db    901 GCCCCCTGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 960

Qy   1097 CCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGG 1156

Db    961 ACTCACTGCCACCTGGGAGCCCTCACTTTCTAACACCCTCGTCACTGGAGTCATCAGTGG 1020

Qy   1157 AATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAG 1216

Db   1021 GATTGCTGTTTGTGTCATCATCTTCTTCATTGGAATTTTGTTCAGAATCTTAAGGAAAAG 1080

Qy   1217 GCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGC 1276

Db   1081 ACAAGCTTCAAGAGGAGCCATGGGGGACTACGTCTTAGGTGAATGTGAGTGACGCACAGC 1140

Qy   1277 CTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATG 1336

Db   1141 CTGCAGACTCCTTGTGGGGAGGACACATAATTGGAGACTCAGAGAGGGTGCAC-TTGCGG 1199

Qy   1337 AGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCC 1395

Db   1200 CATTCTTTGTGTTTCAGGACAGAGTTGAACCTAACAAATGGGAACTGTCCCAGGAGGTCT 1259

Qy   1396 TTGATTTTAGCCTTCTCTGTTCATTTCCTCAAA 1428

Db   1260 CTGCTCTTAGCCTTCCCTGTTCACTCTCAAAAA 1292


RESULT  14
AY007543
LOCUS       AY007543               1320 bp    mRNA    linear   MAM 22-OCT-2000
DEFINITION  Dicerorhinus sumatrensis HFE protein mRNA, complete cds.
ACCESSION   AY007543
VERSION     AY007543.1  GI:10945691
KEYWORDS    .
SOURCE      Sumatran rhinoceros.
  ORGANISM  Dicerorhinus sumatrensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.
REFERENCE   1  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Dicerorhinus sumatrensis"
                     /db_xref="taxon:89632"
     CDS             86. .1132
                     /codon_start=1
                     /product="HFE protein"
                     /protein_id="AAG23703.1"
                     /db_xref="GI:10945692"
                     /translation="MGPRARPALFFLILLRTVAAQGRPPRSHSLRYLFMGASERDHGL
                     PLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWLQLSQSLKGWDHMFIVDF
                     WTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDHLEFCPETLDWRAAES
                     RALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQVPPLVKVTHHVAS
                     AVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQSWVALAVPPGE
                     EQRYTCQVEHPGLDQPLTATWEPSLSNTLVTGVISGIAVCVIIFLIGILFRILRKRQA
                     SRGAMGDYVLAECE"
BASE COUNT      281 a    358 c    398 g    283 t
ORIGIN


  Query Match           62.3%;  Score 897;  DB 4;  Length 1320;
  Best Local Similarity   82.4%;  Pred. No. 4.5e-252;
  Matches 1065;  Conservative    0;  Mismatches  225;  Indels    3;  Gaps    3;

Qy    137 AGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACG 196

Db      2 AGGGTGACTTCTTGGGTCACTCCTATTTTCCCGCCCCACCGAAGCAGCTGACA-CGGGCC 60

Qy    197 GGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCT 256

Db     61 AGCACCCGCGGTCCGAGCCGGGGAAATGGGCCCGCGAGCCCGACCCGCGCTGTTCTTCCT 120

Qy    257 GATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTA 316

Db    121 GATCCTCCTGCGAACCGTGGCCGCGCAGGGGCGACCGCCGCGGTCACACTCTCTGCGCTA 180

Qy    317 CCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTA 376

Db    181 CCTCTTCATGGGCGCCTCAGAGCGGGACCACGGCCTGCCCCTGTTTGAAGCCTTGGGCTA 240

Qy    377 CGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAAC 436

Db    241 TGTGGATGACGAGCTTTTCGTGGCGTACAATCATGAGAGTCGCCGTGCAGAGTCCCGTGC 300

Qy    437 TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA 496

Db    301 CCAGTGGGTCTTGGGTGAGGCCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAA 360

Qy    497 AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAG 556

Db    361 AGGGTGGGATCACATGTTCATCGTCGACTTCTGGACCATCATGGACAACCACAACCACAG 420

Qy    557 CAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTAC 616

Db    421 CAAGGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGGTGCAAGAGGACAACAGCAC 480

Qy    617 CGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACAC 676

Db    481 CAGGGGTTCTGGAAGTACGGGTACGATGGGCAGGACCATCTTGAATTCTGCCCTGAGAC 540

Qy    677 ACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCA 736
```

```
Db    541 ACTGGATTGGAGAGCAGCAGAGTCCAGGGCGCTGACCACCAAGCTGGAGTGGGAAGTGAA 600

Qy    737 CAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCA 796

Db    601 CAAGATTCGGGCCAAGCAGAACAGAGCCTATCTCGAGAGAGACTGCCCTGAGCAGCTGCA 660

Qy    797 GCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGT 856

Db    661 GTGGTTGCTGGAGCTGGGGAGAGGGGTTCTGGACCAGCAAGTGCCTCCTTTGGTGAAGGT 720

Qy    857 GACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCC 916

Db    721 GACTCATCATGTGGCCTCTGCAGTGACCACTCTACGGTGTCAGGCTCTGAACTTCTACCC 780

Qy    917 CCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGA 976

Db    781 CCAGAACATCACCATGAGGTGGTTGAAGGACAGGAAGCCAGTGGATGTCAAGGACGCCGA 840

Qy    977 ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGT 1036

Db    841 GTCTAAGGATGTGCTGCCCAGCGGGGATGGGACCTACCAGAGCTGGGTGGCTTTGGCCGT 900

Qy   1037 ACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 1096

Db    901 GCCCCCTGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 960

Qy   1097 CCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGG 1156

Db    961 CCTCACTGCCACCTGGGAGCCCTCACTTTCTAACACCCTTGTCACTGGAGTCATCAGTGG 1020

Qy   1157 AATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAG 1216

Db   1021 GATTGCTGTTTGTGTCATCATCTTCCTCATTGGAATTTTGTTCAGAATCTTAAGGAAAAG 1080

Qy   1217 GCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGC 1276

Db   1081 ACAAGCTTCGAGAGGAGCCATGGGGGACTACGTCTTAGCTGAATGTGAGTGAAGCACAGC 1140

Qy   1277 CTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATG 1336

Db   1141 CTGCAGACTCCTTGTGGGGAGGACACATAATTGGAGACTCAGAGAGGGTGCAC-TTGTGG 1199

Qy   1337 AGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCC 1395

Db   1200 CGTTCTTCGTGTTTCAGGACAGAGTTGAACCTAACAAATGGGAACTGTCCCAGGAGGTCT 1259

Qy   1396 TTGATTTTAGCCTTCTCTGTTCATTTCCTCAAA 1428

Db   1260 CTGCTCTTAGCCTTCCCTGTTCACTCTCAAAAA 1292


RESULT  15
AY007544
LOCUS       AY007544               1319 bp    mRNA    linear   MAM 22-OCT-2000
DEFINITION  Rhinoceros unicornis HFE protein mRNA, complete cds.
ACCESSION   AY007544
VERSION     AY007544.1  GI:10945693
KEYWORDS    .
SOURCE      greater Indian rhinoceros.
  ORGANISM  Rhinoceros unicornis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
REFERENCE   1  (bases 1 to 1319)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1319)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA

FEATURES             Location/Qualifiers
     source          1. .1319
                     /organism="Rhinoceros unicornis"
                     /db_xref="taxon:9809"
     CDS             86. .1132
                     /codon_start=1
                     /product="HFE protein"
                     /protein_id="AAG23704.1"
                     /db_xref="GI:10945694"
                     /translation="MGPRARPALFFLILLRTVAAQGRPPRSHSLRYLFMGASERDHGL
                     PLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWLQLSQSLKGWDHMFIVDF
                     WTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWMYGYDGQDHLEFCPETLDWRAAES
                     RALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQVPPLVKVTHHVAS
                     AVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQSWVALAVPPGE
                     EQRYTCQVEHPGLDQPLTATWEPSLSNTLVTGVISGIAVCVIIFLIGILFRILRKRQA
                     SRGAMGDYVLAECE"
BASE COUNT      279 a    363 c    397 g    280 t
ORIGIN


  Query Match           62.3%;  Score 896.8;  DB 4;  Length 1319;
  Best Local Similarity   81.9%;  Pred. No. 5.1e-252;
  Matches 1058;  Conservative    0;  Mismatches  232;  Indels    2;  Gaps    2;

Qy    137 AGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACG 196

Db      2 AGGGTGACTTCTTGGGTCACTCCTATTTCCCCGCCCCACCGAAGCAGCTGACA-CGGGCC 60

Qy    197 GGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCT 256

Db     61 AGCACCCGCGGTCCGAGCCGGGGAAATGGGCCCGCGAGCCCGACCCGCGCTCTTCTTCCT 120

Qy    257 GATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTA 316

Db    121 GATCCTCCTGCGGACCGTGGCCGCGCAGGGGCGACCGCCGCGGTCACACTCTCTGCGCTA 180

Qy    317 CCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTA 376

Db    181 CCTCTTCATGGGCGCCTCAGAGCGGGACCATGGCCTGCCCCTGTTTGAAGCCTTGGGCTA 240

Qy    377 CGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAAC 436

Db    241 TGTGGATGATGAGCTTTTCGTGGCCTACAATCACGAGAGTCGCCGTGCAGAGTCCCGTGC 300

Qy    437 TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA 496

Db    301 CCAGTGGGTCTTGGGTGAGGCCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAA 360

Qy    497 AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAG 556

Db    361 AGGGTGGGATCACATGTTCATCGTCGACTTCTGGACCATCATGGACAACCACAACCACAG 420

Qy    557 CAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTAC 616

Db    421 CAAGGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGGTGCAAGAGGACAACAGCAC 480

Qy    617 CGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACAC 676

Db    481 CAGGGGGTTCTGGATGTACGGGTACGACGGGCAGGACCATCTTGAATTCTGCCCTGAGAC 540

Qy    677 ACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCA 736

Db    541 ACTGGATTGGAGAGCAGCAGAGTCCAGGGCCCTGACCACCAAGCTGGAGTGGGAAGTGAA 600

Qy    737 CAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCA 796

Db    601 CAAGATTCGGGCCAAGCAGAACAGAGCCTATCTCGAGAGGGACTGCCCTGAGCAGCTGCA 660

Qy    797 GCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGT 856

Db    661 GTGGTTGCTGGAGCTGGGGAGAGGGGTTCTGGACCAGCAAGTGCCTCCTTTGGTGAAGGT 720

Qy    857 GACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCC 916
```

```
Db    721 GACTCATCATGTGGCCTCTGCAGTGACCACTCTACGGTGTCAGGCTCTGAACTTCTACCC 780

Qy    917 CCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGA 976
          ||||||||||||||| |||| ||||||| | | ||||| |||||| ||||||   |||
Db    781 CCAGAACATCACCATGAGGTGGTTGAAGGACAGGAAGCCAGTGGATGTCAAGGACGCCGA 840

Qy    977 ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGT 1036
            |||| || ||  |||||| ||||||||||||||||||| ||||| |  | ||||||||
Db    841 GTCTAAGGATGTGCTGCCCAGCGGGGATGGGACCTACCAGAGCTGGGTGGCTTTGGCTGT 900

Qy   1037 ACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 1096
          ||||| ||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    901 GCCCCCCGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 960

Qy   1097 CCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGG 1156
          ||||| ||  | |||||||||||||  |||  |||||| |||| ||||||||||||||
Db    961 CCTCACTGCCACCTGGGAGCCCTCACTTTCTAACACCCTCGTCACTGGAGTCATCAGTGG 1020

Qy   1157 AATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAG 1216
           |||||||||| |||| |||||||  ||||||||||||||||| ||| |||||||| ||
Db   1021 GATTGCTGTTTGTGTCATCATCTTCCTCATTGGAATTTTGTTCAGAATCTTAAGGAAAAG 1080

Qy   1217 GCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGC 1276
           || | ||||||||||||||||||| |||||||||||||||| ||||||||  | ||||
Db   1081 ACAAGCTTCAAGAGGAGCCATGGGGGACTACGTCTTAGCTGAATGTGAGTGAAGCACAGC 1140

Qy   1277 CTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATG 1336
          |||||||||  |||||| |||| ||| || | |||||||| |||||| |  | | |    |
Db   1141 CTGCAGACTCCTTGTGGGGAGGACACATAATTGGAGACTCAGAGAGGGTGCAC-TGGTGG 1199

Qy   1337 AGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCT 1396
           | ||||| |||||||||| ||||||||||||||  |  |  |         | |   ||
Db   1200 CGTTCTTCGTGTTTCAGGACAGAGTTGAACCTAACAAATGGGAACTGTCCCAGGAGGTCT 1259

Qy   1397 TGATTTTAGCCTTCTCTGTTCATTTCCTCAAA 1428
             | ||||||||| ||||||| |  |  |||
Db   1260 CTCTCTTAGCCTTCCCTGTTCACTCTCAAAAA 1291


Search completed: June 19, 2002, 08:23:48
Job time: 6314 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:34 ; Search time 2546.54 Seconds
(without alignments)
7632.166 Million cell updates/sec

Title: US-09-497-957-9
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	738.6	51.3	819	10	BG747345	BG747345 602704818
	2	633.8	44.0	1723	11	AK009581	AK009581 Mus muscu
	3	506.8	35.2	570	10	BE272926	BE272926 601171213
c	4	358	24.9	384	10	BF883952	BF883952 PM4-ET020
	5	284.2	19.7	550	10	BI339179	BI339179 364041 MA
	6	276.6	19.2	523	10	BF080089	BF080089 230846 MA
	7	259.4	18.0	489	10	BE994943	BE994943 UI-M-CG0p
	8	251.4	17.5	457	9	AI850020	AI850020 UI-M-BG0-
	9	248.4	17.2	455	10	BE995172	BE995172 UI-M-CG0p
c	10	231.6	16.1	464	9	AA217236	AA217236 mu89b05.r
	11	226.4	15.7	268	10	W21141	W21141 zb51g06.r1
	12	214.6	14.9	546	10	BE809138	BE809138 214520 MA
	13	211.2	14.7	481	9	BB851691	BB851691 BB851691
	14	203.8	14.2	831	10	BI452668	BI452668 603169877
	15	198.8	13.8	502	9	BB858165	BB858165 BB858165
	16	198.2	13.8	392	10	BF465475	BF465475 UI-M-CG0p
	17	187.4	13.0	1208	11	BC005306	BC005306 Homo sapi
	18	174	12.1	394	10	BF464345	BF464345 UI-M-CG0p
c	19	170.4	11.8	444	12	AZ025590	AZ025590 RPCI-23-3
c	20	160	11.1	334	9	AW902003	AW902003 QV0-NN102
c	21	156.8	10.9	536	12	AZ074871	AZ074871 RPCI-23-4
	22	154	10.7	805	10	BG402460	BG402460 602466163
	23	153.6	10.7	1016	9	AL532717	AL532717 AL532717
	24	151.8	10.5	289	10	H33644	H33644 EST109830 R
c	25	151.6	10.5	481	12	AZ025784	AZ025784 RPCI-23-3
	26	146	10.1	531	12	AQ703007	AQ703007 HS_5443_B
	27	145.6	10.1	1031	9	AL513728	AL513728 AL513728
	28	143.6	10.0	1298	11	AK005051	AK005051 Mus muscu
	29	143.2	9.9	1121	9	AL545738	AL545738 AL545738
	30	141.6	9.8	972	9	AL541209	AL541209 AL541209
	31	141.4	9.8	959	9	AL560013	AL560013 AL560013
	32	140.8	9.8	942	9	AL539924	AL539924 AL539924
	33	139.8	9.7	910	9	AL558822	AL558822 AL558822
	34	139.4	9.7	1020	9	AL535151	AL535151 AL535151
	35	138	9.6	1094	10	BM465674	BM465674 AGENCOURT
	36	134.2	9.3	932	9	AL547313	AL547313 AL547313
	37	133.8	9.3	936	10	BI832514	BI832514 603082113
	38	133	9.2	679	10	BG680802	BG680802 602628602
	39	132.6	9.2	937	9	AL540488	AL540488 AL540488
	40	132.6	9.2	1090	10	BM477276	BM477276 AGENCOURT
	41	131.6	9.1	930	9	AL541347	AL541347 AL541347
	42	131.6	9.1	979	9	AL542403	AL542403 AL542403
	43	131.6	9.1	979	9	AL557373	AL557373 AL557373
	44	131.4	9.1	538	9	BE232472	BE232472 137492 MA
	45	131.2	9.1	916	10	BG744157	BG744157 602722954

ALIGNMENTS

RESULT 1
BG747345
LOCUS BG747345 819 bp mRNA linear EST 15-MAY-2001
DEFINITION 602704818F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4857941 5', mRNA sequence.
ACCESSION BG747345
VERSION BG747345.1 GI:14057998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 819)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1711 row: d column: 06
High quality sequence stop: 792.
FEATURES Location/Qualifiers
source 1. .819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4857941"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 202 a 201 c 235 g 181 t
ORIGIN

Query Match 51.3%; Score 738.6; DB 10; Length 819;
Best Local Similarity 96.1%; Pred. No. 2.2e-194;
Matches 789; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

```
Qy    384 GACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGG 443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGG 60

Qy    444 GTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG 503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG 120

Qy    504 GATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAG 563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAG 180

Qy    564 TCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGC 623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGC 240

Qy    624 TACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGAT 683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGAT 300

Qy    684 TGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATT 743
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    301 TGGAGAGCAGCAGAACCCA-GGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATT 359

Qy    744 CGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTG 803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTG 419

Qy    804 CTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACAT 863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 CTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACAT 479

Qy    864 CATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAAC 923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 CATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAAC 539

Qy    924 ATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAA 983
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    540 ATCACCATGAAGTGGCTGAAGGATAAGCAG-CAATGGATGCCAAGGAGTTCGAACCTAAA 598

Qy    984 GACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCT 1043
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    599 GACGTATTGCCCAATGGGGATGGGACCTACCA-GGCTGGATAACCTTGGCTGTACCCCCT 657

Qy   1044 GGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATT 1103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 GGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATT 717

Qy   1104 GTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCT 1163
          ||||||| ||   ||    |||||||||||       | | |||||||    | | |
Db    718 GTGATCTTGGGAGCCCTCACGTCTGGCACCTAGTCATTGGAAGTCATCCAGTGGAATTGC 777

Qy   1164 GTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAAT 1204
            |||||||||||||||||| ||||||||||||||||||||
Db    778 TGTTTTGTCGTCATCTTGTTCCTTGGAATTTTGTTCATAAT 818
```

RESULT 2
AK009581
LOCUS AK009581 1723 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310032M04:hemochromatosis, full insert sequence.

ACCESSION AK009581
VERSION AK009581.1 GI:12844462
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:2310032M04.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1723)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase

```
          and subsequently enriched for full-length by cap-trapper. cDNA went
          through one round of normalization to Rot = 5.0 and subtraction to
          Rot = 25.0. Second strand cDNA was prepared with the primer adapter
          of sequence [5'
          GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
          with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
          Host: SOLR.
FEATURES             Location/Qualifiers
     source          1. .1723
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="MGD:MGI:1905246"
                     /db_xref="taxon:10090"
                     /clone="2310032M04"
                     /sex="male"
                     /tissue_type="tongue"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     gene            99. .1178
                     /gene="Hfe"
     CDS             99. .1178
                     /gene="Hfe"
                     /note="data source:MGD, source key:MGI:109191,
                     evidence:ISS
                     hemochromatosis
                     putative"
                     /codon_start=1
                     /protein_id="BAB26373.1"
                     /db_xref="GI:12844463"
                     /translation="MSLSAGLPVRPLLLLLLLWSVAPQALPPRSHSLRYLFMGASEP
                     DLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWLHLSQSLKGWDYMF
                     IVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWRYGYDGQDHLEF
                     CPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLELGRGVLGQQV
                     PTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVLPNGDETY
                     QGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWEPLQSQAMIIGIISGVTVCAIFLVG
                     ILFLILRKRKASGGTMGGYVLTDCE"
     polyA_signal    1695. .1700
                     /note="putative"
     polyA_site      1723
                     /note="putative"
BASE COUNT      406 a    456 c    454 g    407 t
ORIGIN
```

```
  Query Match           44.0%;  Score 633.8;  DB 11;  Length 1723;
  Best Local Similarity   70.2%;  Pred. No. 4.6e-165;
  Matches  956;  Conservative    0;  Mismatches  357;  Indels   48;  Gaps    6;

Qy    105 CTGGGCATCTCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTT 164
          | ||  |  ||  || |||  |||||  |||  ||||||||||||  |   | ||| |||||
Db      1 CCGGAGAGGTCTGGAACCTCAGCAATGGCTACAGGGTGACTTCTTGGATCCTCCACGTTT 60

Qy    165 CCCCGCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATG 224
          ||    ||     ||||    |  ||   | |        || || ||           |
Db     61 CCAGATCCTAGTGAAGACCGGTGGACCCAGC-------TGAGGACATGAGCCTATCAGCT 113

Qy    225 GGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAG 284
          || |  |  |   |||||  ||| || || ||| | ||  ||  | ||| || || ||||
Db    114 GGGCTCCCTGTGCGGCCGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAG 173

Qy    285 GGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGAC 344
          |  |     | |||||||||| |||||    ||||||||||||||||||||||||  |||
Db    174 GCACTGCCACCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGAC 233

Qy    345 CTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTAT 404
          || || ||  | |||||||| |||  |||||| ||||||||||||||| || |||| |||
Db    234 CTCGGGCTGCCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTAC 293

Qy    405 GATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGC 464
          ||||||||||||||||||  ||||||  |  | || ||| | |        ||  ||||||
Db    294 AATCATGAGAGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGC 353

Qy    465 CAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGAC 524
```

```
          ||| |||||||||| |||||||||||| |||||||||||||  |||||||||  || |||
Db    354 CAGCTGTGGCTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGAC 413

Qy    525 TTCTGGACTATTATGGAAAATCACAACCACAGCAAG---------------------- 560
          |||||||| || ||||  ||  | ||||||| |||
Db    414 TTCTGGACCATCATGGGCAACTATAACCACAGTAAGGTCACGAAGTTGGGAGTGGTGTCC 473

Qy    561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
          |||||||||| |||||||||  |||| ||||||||  |||| ||||||||||||||||
Db    474 GAGTCCCACATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGC 533

Qy    621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
          ||||  |||||  || || ||||| ||||| || ||||| ||||||||||  | |||||
Db    534 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTA 593

Qy    681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
          | |||||  ||||| || ||  |||||||| |||||||| |||| |||||   |||||||
Db    594 AACTGGAGCGCAGCCGAGCCAGGGGCCTGGGCCACCAAGGTGGAATGGGACGAGCACAAG 653

Qy    741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
          || || ||||  |||||||||| |||||||||| |||||||| |  |||||| | | |
Db    654 ATCCGTGCCAAACAGAACAGGGACTACCTGGAGAAGGACTGCCCCGAGCAGCTGAAACGG 713

Qy    801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACA 860
          | |||||||||||||||| ||| |||  || ||||||||| |||||||||| |||||
Db    714 CTCCTGGAGCTGGGGAGAGGCGTTCTGGGACAGCAAGTGCCTACTTTGGTGAAAGTGACT 773

Qy    861 CATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG 920
          |  |-|   | ||||| | | |||| ||||| |||||| |||  || ||| || |||||||
Db    774 CGCCACTGGGCCTCTACGGGGACCTCTCTAAGGTGTCAGGCTCTGGACTTCTTCCCCCAG 833

Qy    921 AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCT 980
          |||||||| |||| |||||||| || || ||| |||||||||| ||  || | ||
Db    834 AACATCACTATGAGGTGGTTGAAGGACAACCAACCACTGGATGCCAAAGATGTCAACCCC 893

Qy    981 AAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCC 1040
          |  | ||  | || || |||||||  ||||||  || ||||||  | || |||||  ||  ||
Db    894 GAGAAGGTGCTACCTAACGGGGATGAGACCTATCAAGGCTGGCTGACATTGGCCGTGGCC 953

Qy   1041 CCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTC 1100
          |||||||| |||   || |  || || || ||||||||||||||||||||| ||||| |||
Db    954 CCTGGGGACGAGACAAGGTTCACCTGTCAAGTGGAGCACCCAGGCCTGGACCAGCCTCTC 1013

Qy   1101 ATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATT 1160
          | ||     |||||||||| |  |||    || |  | || ||| |||||||||| |
Db    1014 ACTGCCTCTTGGGAGCCCTTGCAATCTCAGGCCATGATTATCGGAATCATCAGTGGAGTC 1073

Qy   1161 GCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAG 1220
          | || | || | || |||||    |||||||| ||||| ||| |||||||| ||| ||
Db    1074 ACCGTCTGTGCCATCTTCTTGG---TTGGAATTCTGTTCCTAATCTTAAGGAAAAGGAAG 1130

Qy   1221 GGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGC 1280
          | |||| ||||| |||||||   ||| ||||||  | ||  ||||||||   |||||||||
Db    1131 GCTTCAGGAGGAACCATGGGTGGCTATGTCTTAACAGACTGTGAGTGATCTGCAGCCTGC 1190

Qy   1281 AGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCT 1340
          ||  ||| |  |  || | ||  | | | ||||  || |||    | |     |
Db    1191 TGAACCACGGAAGAGAGAAAACTCAGCCAAAGACTTGGAGGGGGCACACTTGCTCCA--- 1247

Qy   1341 CTTCATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCCTTGA 1399
                |  |||| | ||||| |||||| ||| ||||| |||||||  |   |  ||
Db    1248 -------CTGTAGGACACAGTTGGACCTAACACACAGAAACTGCCTGAGAA---CTGTGC 1297

Qy   1400 TTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          | ||||| |||||||||| || || ||   | |||| |||
Db    1298 TCTTAGCTTTCTCTGTTCACTTTCTTAAGGTGTTTTCTCCA 1338
```

```
RESULT   3
BE272926
LOCUS       BE272926                570 bp    mRNA    linear   EST 13-JUL-2000
DEFINITION  601171213F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:3544803 5',
            mRNA sequence.
```

```
ACCESSION   BE272926
VERSION     BE272926.1  GI:9147279
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM240  row: j  column: 04
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3544803"
                     /clone_lib="NIH_MGC_14"
                     /tissue_type="renal cell adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      140 a    148 c    175 g    107 t
ORIGIN
```

Query Match 35.2%; Score 506.8; DB 10; Length 570;
Best Local Similarity 99.4%; Pred. No. 6.3e-130;
Matches 519; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy     560 GGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGA 619
           | ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      48 GCAGTCCCACACCGTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGA 107

Qy     620 GGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACT 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108 GGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACT 167

Qy     680 GGATTGGAGAGCAGCAGAACCCAG-GGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACA 738
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     168 GGATTGGAGAGCAGCAGAACCCAGTGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACA 227

Qy     739 AGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGC 798
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 AGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGC 287

Qy     799 AGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGA 858
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 AGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGA 347

Qy     859 CACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCC 918
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     348 CACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCC 407

Qy     919 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAAC 978
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     408 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAAC 467

Qy     979 CTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTAC 1038
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     468 CTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTAC 527

Qy    1039 CCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           |||||||||||||||||||||||||||||||||||||||||||
Db     528 CCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 569
```

```
RESULT   4
BF883952/c
LOCUS       BF883952                384 bp    mRNA    linear   EST 17-JAN-2001
DEFINITION  PM4-ET0209-151200-003-f07 ET0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF883952
VERSION     BF883952.1  GI:12274078
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 384)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM4&t2=PM4-ET0209-
            151200-003-f07&t3=2000-12-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 384.
FEATURES             Location/Qualifiers
     source          1. .384
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ET0209"
                     /dev_stage="Adult"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       92 a    112 c     87 g     93 t
ORIGIN
```

Query Match 24.9%; Score 358; DB 10; Length 384;
Best Local Similarity 99.7%; Pred. No. 1.4e-88;
Matches 369; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy     325 TGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATG 384
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384 TGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATG 325

Qy     385 ACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGG 444
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    324 ACCAGCTG-TCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGG 266

Qy    445 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGG 504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGG 206

Qy    505 ATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGT 564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 ATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGT 146

Qy    565 CCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCT 624
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 CCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCT 86

Qy    625 ACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATT 684
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 ACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATT 26

Qy    685 GGAGAGCAGC 694
          ||||||||||
Db     25 GGAGAGCAGC 16
```

RESULT 5
BI339179
LOCUS BI339179 550 bp mRNA linear EST 30-JUL-2001
DEFINITION 364041 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BI339179
VERSION BI339179.1 GI:15032462
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 550)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 100 row: C column: 24
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .550
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 108 a 180 c 164 g 98 t
ORIGIN

Query Match 19.7%; Score 284.2; DB 10; Length 550;
Best Local Similarity 74.1%; Pred. No. 5.8e-68;
Matches 413; Conservative 0; Mismatches 113; Indels 31; Gaps 3;

```
Qy    140 GTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACGGGG 199
          |||||||||  |||||    |   ||  |||  |||  ||||   |  |   ||       |  ||
Db      1 GTGACTTCTTGAGCCCCTCGGTTTCCCCGCCCCACCAACGGCCGAAAAG-----CCCTGG 55

Qy    200 ACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGAT 259
          | | ||||  |  ||||  |||   ||  ||||||  |  ||||  ||||||||||||||  ||||||||
Db     56 AAGCGCGGTCCGAGCCGGG--AAAGGGCCCACAAGCCCGGCCGGCGCTTCTGCTCCTGAT 113

Qy    260 GCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCT 319
          ||  |||  |||||  ||  |  |||||  |||   ||  |||  |||||||  ||||  ||  |||
Db    114 CCTCCTGCGGACCGTGGCCACGCAGGAGCGAGCGCCGCGGCCACACTCCCTGCTCTTCCT 173

Qy    320 CTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGT 379
          |||||||||  |||||  ||||   ||  ||  ||  ||   ||  |||||||  ||  ||||||||||
Db    174 CTTCATGGGCGCCTCGGAGCCAGATCTCGGGCTGCCCCTGTTTGAGGCCTTGGGCTACGT 233

Qy    380 GGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCC 439
          |||  ||||||||||||  ||||  |||   ||||  ||||||||||||||   |||||  ||   |  ||
Db    234 GGACGACCAGCTGTTTGTGTCCTACAATCACGAGAGTCGCCGTGCAGAGCCTCGCGCCCC 293

Qy    440 ATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGG 499
          |||  |  ||  |||      ||  |  |||  ||||||||||||||  ||  |||||  ||||||||
Db    294 CTGGCTCTCGGGTAAGGCCTCCAACCAACTGTGGCTGCAGCTAAGCCAGAGCCTGAAAGG 353

Qy    500 GTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAA 559
          ||||||||||||||||||  ||  ||||||||||  ||  |||||  ||  ||||||  |||||||
Db    354 GTGGGATCACATGTTCACCGTGGACTTCTGGACCATCATGGACAACCACAACTACAGCAA 413

Qy    560 G----------------------GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGA 595
          |                      ||||||||||||||||||  ||||||||||||||
Db    414 GATAACGAAGCTGGGAGTGTTGCCAGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGA 473

Qy    596 AATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCA 655
          |  ||||||  ||||||||  |||   ||  |  |||||||||  |||||  ||||||||||||||
Db    474 AGTGCAAGCGGACAACAGCACCAGAGGGTTCTGGAAGTATGGGTACGATGGGCAGGACCA 533

Qy    656 CCTTGAATTCTGCCCTG 672
          |||  ||  |||  |||||
Db    534 CCTGGAGTTCCACCCTG 550
```

RESULT 6
BF080089
LOCUS BF080089 523 bp mRNA linear EST 18-OCT-2000
DEFINITION 230846 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF080089
VERSION BF080089.1 GI:10873919
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 523)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 48 row: E column: 9
Seq primer: ATTTAGGTGACACTATAG.

```
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 2PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT      103 a    175 c    152 g     93 t
ORIGIN
```

```
  Query Match             19.2%;  Score 276.6;  DB 10;  Length 523;
  Best Local Similarity   74.2%;  Pred. No. 7.4e-66;
  Matches  388;  Conservative    0;  Mismatches  109;  Indels   26;  Gaps    2;

Qy    119 AGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAA 178
          ||||| |||||| ||| |||||||||||| |||||| | | ||||||||||||||    |
Db      2 AGCCTTGGCAATGGCTCCAGGGTGACTTCTTGAGCCACCTCGGTTTCCCCGCCCCAGCCA 61

Qy    179 AGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAG 238
          ||  || || |   | | || |     | ||    |  ||||||||||||| | |||| |
Db     62 AGCGGCCGAAAAGCACCTGGAAGCGCGGTCCGAGCC--GGGAAATGGGCCCACAAGCCCG 119

Qy    239 GCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCG 298
          |||||||||||| |||||||| ||  ||| ||||| || |  ||| |  ||   || |||
Db    120 GCCGGCGCTTCTGCTCCTGATCCTCCTGCGGACCGTGGCCACGCATGATCGAGCGCCGCG 179

Qy    299 TTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTT 358
            ||||||| |||| || |||||||||||| ||||| ||||  || || ||  ||  | |
Db    180 GCCACACTCCCTGCTCTTCCTCTTCATGGGCGCCTCGGAGCCAGATCTCGGGCTGCCCCT 239

Qy    359 GTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCG 418
          ||||||  || |||||||||||||| |||||||||| |||| |||  |||| ||||||||
Db    240 GTTTGAGGCCTTGGGCTACGTGGACGACCAGCTGTTTGTGTCCTACAATCACGAGAGTCG 299

Qy    419 CCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCA 478
          |||||  ||||| ||  | || ||| | ||  |||     || | |||  ||||||||||
Db    300 CCGTGCAGAGCCTCGCGCCCCCTGGCTCTCGGGTAAGGCCTCCAACCAACTGTGGCTGCA 359

Qy    479 GCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTAT 538
          ||| || ||||| |||||||||||||||||||||||||| || ||||||||||| || ||
Db    360 GCTAAGCCAGAGCCTGAAAGGGTGGGATCACATGTTCACCGTGGACTTCTGGACCATCAT 419

Qy    539 GGAAAATCACAACCACAGCAAG-----------------------GAGTCCCACACCCT 574
          ||| || |||||| ||||||||                       ||||||||||||||
Db    420 GGACAACCACAACTACAGCAAGATAACGAAGCTGGGAGTGTGGCCAGAGTCCCACACCCT 479

Qy    575 GCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACC 617
          |||||| |||||||||||||| ||||||  ||||||| |||
Db    480 GCAGGTGATCCTGGGCTGTGAAGTGCAAGCGGACAACAGCACC 522
```

```
RESULT   7
BE994943
LOCUS       BE994943                489 bp    mRNA    linear   EST 05-OCT-2000
DEFINITION  UI-M-CG0p-bik-d-03-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
            UI-M-CG0p-bik-d-03-0-UI 3', mRNA sequence.
ACCESSION   BE994943
VERSION     BE994943.1  GI:10678689
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
```

```
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
            clones from RESEARCH GENETICS. It should be noted that Bento Soares
            is generating a small number of additional specialized
            non-redundant arrays of BMAP cDNAs whose availability will be
            considered under appropriate and limited collaborative arrangements
            The following repetitive elements were found in this cDNA sequence:
            1-31, >(CAG)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-CG0p-bik-d-03-0-UI"
                     /clone_lib="NIH_BMAP_Ret4_S2"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_Ret4_S2 library is a subtracted library,
                     ultimately derived from mouse retina tissue libraries at
                     various stages of development. For a detailed description
                     of the library from which this clone was derived,  please
                     visit our web site at brainest.eng.uiowa.edu.
                     TAG_SEQ=None found"
BASE COUNT      103 a    133 c    144 g    109 t
ORIGIN
```

```
  Query Match             18.0%;  Score 259.4;  DB 10;  Length 489;
  Best Local Similarity   73.3%;  Pred. No. 4.4e-61;
  Matches  357;  Conservative    0;  Mismatches  106;  Indels   24;  Gaps    1;

Qy    238 GGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGC 297
          |||||  ||| || || ||| | ||  ||  | ||| || || |||||  |     | ||
Db      3 GGCCGCTGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAGGCACTGCCACCGC 62

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          |||||||  |||||    ||||||||||||||||||||||||  |||||  || ||  | |
Db     63 GTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGACCTCGGGCTGCCTT 122

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
          |||||||  |||  ||||||  |||||||||||||||  ||  ||||  |||  ||||||||||
Db    123 TGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATCATGAGAGTC 182

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||  ||||||  |  | || |||  | |        ||  |||||||| |||||||||
Db    183 GCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGCTGTGGCTGC 242

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          | ||||||||||  ||||||||||||||  |||||||||  || ||||||||||| || |
Db    243 ATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTCTGGACCATCA 302

Qy    538 TGGAAAATCACAACCACAGCAAG-----------------------GAGTCCCACACCC 573
          |||  ||  | ||||||||| |||                       |||||||||| ||
Db    303 TGGGCAACTATAACCACAGTAAGGTCACGAAGTTGGGAGTGGTGTCCGAGTCCCACATCC 362

Qy    574 TGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGT 633
          |||||||  |||| ||||||||  |||| ||||||||||||||   |||| |||||  |
Db    363 TGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGCGGCTTCTGGAGAT 422

Qy    634 ACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAG 693
          | || ||||| ||||| || ||||| ||||||||||  | |||||  | ||||| ||||
Db    423 ATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTAAACTGGAGCGCAG 482
```

```
Qy     694 CAGAACC 700
           | || ||
Db     483 CCGAGCC 489


RESULT   8
AI850020
LOCUS       AI850020                457 bp    mRNA    linear   EST 15-JUL-1999
DEFINITION  UI-M-BG0-aib-g-10-0-UI.s1 NIH_BMAP_MSC Mus musculus cDNA clone
            UI-M-BG0-aib-g-10-0-UI 3', mRNA sequence.
ACCESSION   AI850020
VERSION     AI850020.1  GI:5493926
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
            available by the means that is soon to be determined. When NIH
            determines the means for distribution of the BMAP cDNA clones, this
            record will be updated accordingly when that means is determined.
            The following repetitive elements were found in this cDNA sequence:
            3-30, >(CAG)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BG0-aib-g-10-0-UI"
                     /clone_lib="NIH_BMAP_MSC"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_MSC library is a non-normalized library
                     constructed from  mouse spinal cord. The tag is a string
                     of 5 nucleotides present  between the Not I site and the
                     oligo-dT track.  The library was  constructed as described
                     by Bonaldo, Lennon and Soares,  Genome Research 6: 791-806
                     , 1996. Tissue provided by Ms. Annie Novakovich,
                     Zivic-Miller Laboratories.
                     TAG_LIB=NIH_BMAP_MSC
                     TAG_TISSUE=spinal-cord
                     TAG_SEQ=TCGAA"
BASE COUNT      100 a    129 c    144 g     84 t
ORIGIN


  Query Match            17.5%;  Score 251.4;  DB 9;  Length 457;
  Best Local Similarity  79.6%;  Pred. No. 7.2e-59;
  Matches  297;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy     561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
           |||||||||| |||||||||  |||| ||||||||  |||| ||||||||||||||||
Db      85 GAGTCCCACATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGC 144

Qy     621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
           |||| |||||  || || ||||| ||||| || ||||| |||||||||||  | |||||
Db     145 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTA 204

Qy     681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
           | ||||| ||||| || ||  |||||||| |||||||| |||| |||||   |||||||
Db     205 AACTGGAGCGCAGCCGAGCCAGGGGCCTGGGCCACCAAGGTGGAATGGGACGAGCACAAG 264

Qy     741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
           || || ||||  |||||||||| |||||||||| ||||||||| |  |||||| | | |
Db     265 ATCCGTGCCAAACAGAACAGGGACTACCTGGAGAAGGACTGCCCCGAGCAGCTGAAACGG 324

Qy     801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACA 860
           | |||||||||||||||| ||| |||  || ||||||||| |||||||||| |||||
Db     325 CTCCTGGAGCTGGGGAGAGGCGTTCTGGGACAGCAAGTGCCTACTTTGGTGAAAGTGACT 384

Qy     861 CATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG 920
           |  ||   | ||||| | | |||| |||| |||||| |||  || ||| || |||||||
Db     385 CGCCACTGGGCCTCTACGGGGACCTCTCTAAGGTGTCAGGCTCTGGACTTCTTCCCCCAG 444

Qy     921 AACATCACCATGA 933
           ||||||| ||||
Db     445 AACATCACTATGA 457


RESULT   9
BE995172
LOCUS       BE995172                455 bp    mRNA    linear   EST 05-OCT-2000
DEFINITION  UI-M-CG0p-bil-h-10-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
            UI-M-CG0p-bil-h-10-0-UI 3', mRNA sequence.
ACCESSION   BE995172
VERSION     BE995172.1  GI:10679153
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
            clones from RESEARCH GENETICS. It should be noted that Bento Soares
            is generating a small number of additional specialized
            non-redundant arrays of BMAP cDNAs whose availability will be
            considered under appropriate and limited collaborative arrangements
            The following repetitive elements were found in this cDNA sequence:
            1-31, >(CAG)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-CG0p-bil-h-10-0-UI"
                     /clone_lib="NIH_BMAP_Ret4_S2"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_Ret4_S2 library is a subtracted library,
```

ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu.
TAG_SEQ=None found"
BASE COUNT 99 a 129 c 144 g 83 t
ORIGIN

Query Match 17.2%; Score 248.4; DB 10; Length 455;
Best Local Similarity 79.5%; Pred. No. 4.9e-58;
Matches 294; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```
Qy    561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
          |||||||||| |||||||||  |||| ||||||||  |||| ||||||||||||||||
Db     86 GAGTCCCACATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGC 145

Qy    621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
          |||| |||||  || || ||||| ||||| || ||||| |||||||||||  | |||||
Db    146 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTA 205

Qy    681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
          | ||||| ||||| || ||  |||||||| |||||||| |||| |||||   |||||||
Db    206 AACTGGAGCGCAGCCGAGCCAGGGGCCTGGGCCACCAAGGTGGAATGGGACGAGCACAAG 265

Qy    741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
          || || ||||  |||||||||| |||||||||| |||||||| |  |||||| | | |
Db    266 ATCCGTGCCAAACAGAACAGGGACTACCTGGAGAAGGACTGCCCCGAGCAGCTGAAACGG 325

Qy    801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACA 860
          | |||||||||||||||| ||| |||  || |||||||| |||||||||| |||||
Db    326 CTCCTGGAGCTGGGGAGAGGCGTTCTGGGACAGCAAGTGCCTACTTTGGTGAAAGTGACT 385

Qy    861 CATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG 920
          |  ||   | ||||| | | |||| ||||| |||||| |||  || ||| || |||||||
Db    386 CGCCACTGGGCCTCTACGGGGACCTCTCTAAGGTGTCAGGCTCTGGACTTCTTCCCCCAG 445

Qy    921 AACATCACCA 930
          ||||||||  |
Db    446 AACATCACTA 455
```

RESULT 10
AA217236/c
LOCUS AA217236 464 bp mRNA linear EST 06-FEB-1997
DEFINITION mu89b05.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone IMAGE:652689 5' similar to TR:G940354 G940354 CLASS I HISTOCOMPATIBILITY ANTIGEN-LIKE PROTEIN. ;, mRNA sequence.
ACCESSION AA217236
VERSION AA217236.1 GI:1826237
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 464)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:398537

Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .464
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:652689"
/clone_lib="Soares mouse lymph node NbMLN"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 101 a 136 c 119 g 108 t
ORIGIN

Query Match 16.1%; Score 231.6; DB 9; Length 464;
Best Local Similarity 74.3%; Pred. No. 2.4e-53;
Matches 330; Conservative 0; Mismatches 89; Indels 25; Gaps 2;

```
Qy    294 CTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTT 353
          | |||||||| |||||    ||||||||||||||||||||||||  ||||| || ||
Db    453 CCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGACCTCGGGCTG 394

Qy    354 TCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAG 413
          | |||||||| |||  |||||| ||||||||||||| || |||| |||  |||||||
Db    393 CCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATCATGAG 334

Qy    414 AGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGG 473
          ||||||||||  |||||| |  | || ||| | |       ||  |||||||| |||||
Db    333 AGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGCTGTGG 274

Qy    474 CTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACT 533
          ||||| |||||||||| ||||||||||||||  |||||||||  || ||||||||||
Db    273 CTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTCTGGACC 214

Qy    534 ATTATGGAAAATCACAACCACAGCAAG----------------------GAGTCCCAC 569
          || ||||  ||  | ||||||||  |||                      |||||||||
Db    213 ATCATGGGCAACTATAACCACAGTAAGGTCACGAAGTTGGGAGTGGTGTCCGAGTCCCAC 154

Qy    570 ACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGG 629
          | ||||||||| |||| |||||||| |||| ||||||||||||||   |||| ||||
Db    153 ATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGCGGCTTCTGG 94

Qy    630 AAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGA 689
          |  || || ||||| ||||| || ||||| ||||||||||  | |||||  | |||||
Db     93 AGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTAAACTGGAG- 33

Qy    690 GCAGCAGAACCCAGGGCCTGGCCC 713
          ||||| || ||  || |  ||| |
Db     34 GCAGCCGAGCCAGGGCCTGGGCAC 11
```

RESULT 11
W21141
LOCUS W21141 268 bp mRNA linear EST 20-AUG-1996
DEFINITION zb51g06.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:307162 5', mRNA sequence.
ACCESSION W21141
VERSION W21141.1 GI:1298018
KEYWORDS EST.
SOURCE human.

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 268)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1642   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 173.
FEATURES             Location/Qualifiers
     source          1. .268
                     /organism="Homo sapiens"
                     /db_xref="GDB:1250574"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:307162"
                     /clone_lib="Soares_fetal_lung_NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library, Soares
                     fetal heart NbHH19W."
BASE COUNT       79 a     55 c     64 g     69 t      1 others
ORIGIN
```

```
  Query Match           15.7%;  Score 226.4;  DB 10;  Length 268;
  Best Local Similarity   99.2%;  Pred. No. 5.1e-52;
  Matches  238;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      1 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG-ACTACGTCTTAGCTGAAC 59

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 119

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 179

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    180 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAAATTTCCCCA 239
```

```
RESULT  12
BE809138
LOCUS       BE809138                546 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  214520 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE809138
VERSION     BE809138.1  GI:10240250
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 71  row: A  column: 5
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT      126 a    139 c    180 g    100 t      1 others
ORIGIN
```

```
  Query Match           14.9%;  Score 214.6;  DB 10;  Length 546;
  Best Local Similarity   85.9%;  Pred. No. 1.4e-48;
  Matches  238;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy    561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
          |||||||||||||||||| |||||||||||||  ||||||| |||||||| |||
Db    139 GAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGTTGCAAGAGGACAACAGCACCAGA 198

Qy    621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
          || | |||||||||||||| |||||||||||||| ||||||||| | ||||| ||||||
Db    199 GGGTTCTGGAAGTACGGGTACGATGGGCAGGACCATCTTGAATTCCGGCCTGAGACACTG 258

Qy    681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
          |||||||||||||||| |||||||||     ||||||||||||||||||  | |||||
Db    259 GATTGGAGAGCAGCAGAGCCCAGGGCCCAAGTCACCAAGCTGGAGTGGGAAGTGAACAAG 318

Qy    741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
          |||||||||| |||||||||||||||||||  |||| ||||| |  ||||||| |||
Db    319 ATTCGGGCCAAGCAGAACAGGGCCTACCTGGATCGGGATTGCCCCGAGCAGCTGCTGCAT 378

Qy    801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          |||||||||||||||||||   | |||| || ||||
Db    379 TTGCTGGAGCTGGGGAGAGGGCCTCTGGAGCAGCAAG 415
```

```
RESULT  13
BB851691
LOCUS       BB851691                481 bp    mRNA    linear   EST 26-NOV-2001
DEFINITION  BB851691 RIKEN full-length enriched, B16 F10Y cells Mus musculus
            cDNA clone G370002P09 5', mRNA sequence.
ACCESSION   BB851691
VERSION     BB851691.1  GI:17093145
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .481
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G370002P09"
                     /clone_lib="RIKEN full-length enriched, B16 F10Y cells"
                     /cell_type="B16 F10Y cells"
BASE COUNT       95 a    134 c    140 g    112 t
ORIGIN


  Query Match           14.7%;  Score 211.2;  DB 9;  Length 481;
  Best Local Similarity   68.8%;  Pred. No. 1.1e-47;
  Matches  308;  Conservative    0;  Mismatches  133;  Indels    7;  Gaps    1;

Qy    114 TCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCC 173
          .   ||  || |||  ||||| |||  |||||||||||| |   | |||  |||||||    ||
Db      9 TCTGGAACCTCAGCAATGGCTACAGGGTGACTTCTTGGATCCTCCACGTTTCCAGATCCT 68

Qy    174 CCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGA 233
              ||||   |  ||   | |       || || ||          |   || | |
Db     69 AGTGAAGACCGGTGGACCCAGC-------TGAGGACATGAGCCTATCAGCTGGGCTCCCT 121

Qy    234 GCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTG 293
          |   |||||  ||| || || ||| | ||  ||  | ||| || || |||||  |
Db    122 GTGCGGCCGCTGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAGGCACTGCCA 181

Qy    294 CTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTT 353
          | ||||||||| |||||    ||||||||||||||||||||||||  ||||| || ||
Db    182 CCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGACCTCGGGCTG 241

Qy    354 TCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAG 413
          | |||||||| |||  |||||| |||||||||||||| || |||| |||  ||||||||
Db    242 CCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATCATGAG 301

Qy    414 AGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGG 473
          ||||||||||  |||||| |  | || ||| | |       ||  |||||||| ||||
Db    302 AGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGCTGTGG 361

Qy    474 CTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACT 533
          |||||  |||||||||||  ||||||||||||||  |||||||||  || ||||||||||
Db    362 CTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTCTGGACC 421

Qy    534 ATTATGGAAAATCACAACCACAGCAAGG 561
          || ||||  ||  | ||||||||| ||||
Db    422 ATCATGGGCAACTATAACCACAGTAAGG 449


RESULT  14
BI452668
LOCUS       BI452668                831 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603169877F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5249395 5',
            mRNA sequence.
ACCESSION   BI452668
VERSION     BI452668.1  GI:15243324
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11629  row: j  column: 20
            High quality sequence stop: 818.
FEATURES             Location/Qualifiers
     source          1. .831
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5249395"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT      207 a    220 c    189 g    215 t
ORIGIN
```

```
  Query Match             14.2%;  Score 203.8;  DB 10;  Length 831;
  Best Local Similarity   68.5%;  Pred. No. 1.7e-45;
  Matches  346;  Conservative    0;  Mismatches  142;  Indels   17;  Gaps    4;

Qy    935 GTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCC 994
          |||| ||||||| || || ||| |||||||||| ||  || | ||  |  | ||  | ||
Db      2 GTGGTTGAAGGACAACCAACCACTGGATGCCAAAGATGTCAACCCCGAGAAGGTGCTACC 61

Qy    995 CAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCA 1054
          || ||||||| |||||| || |||||| | || ||||| ||  |||||||||| |||
Db     62 TAACGGGGATGAGACCTATCAAGGCTGGCTGACATTGGCCGTGGCCCCTGGGGACGAGAC 121

Qy   1055 GAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGA 1114
          || |  || || || |||||||||||||||||||| ||||| |||| ||     |||||
Db    122 AAGGTTCACCTGTCAAGTGGAGCACCCAGGCCTGGACCAGCCTCTCACTGCCTCTTGGGA 181

Qy   1115 GCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGT 1174
          |||||  |  |||    || |  | |||||| ||||||||||  |  |  | | |   |
Db    182 GCCCTTGCAATCTCAGGCCATGATTATTGGAATCATCAGTGGGGTCACCATCTGT---GC 238

Qy   1175 CATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGC 1234
          |||||| ||  |||||||| ||||| |||| |||||||| ||| ||| |||| ||||| |
Db    239 CATCTTCTTAGTTGGAATTCTGTTCCTAATCTTAAGGAAAAGGAAGGCTTCAGGAGGAAC 298

Qy   1235 CATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGG 1294
          ||||||   ||| |||||| | ||  |||||||   |||||||| ||  ||| |  |
Db    299 CATGGGTGGCTATGTCTTAACAGACTGTGAGTGATCTGCAGCCTGCTGAACCACGGAAGA 358

Qy   1295 AAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGG 1354
          || | ||  | | | |||||   || |||    | |     |          |  |||
Db    359 GAGAAAACTCAGCCAAAGACTTGGAGGGGGCACACTTGCTCCA----------CTGTAGG 408

Qy   1355 AGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCT 1413
          | | ||||| |||||| ||| ||||| |||||||  |   |  || | ||||||||||||
Db    409 ACACAGTTGGACCTAACACACAGAAACTGCCTGAGAA---CTGTGCTCTTAGCCTTCTCT 465

Qy   1414 GTTCATTTCCTCAAAAAGATTTCCC 1438
          ||||| || || ||   ||||  ||
Db    466 GTTCACTTTCTTAAGGTGATTCTCC 490
```

```
RESULT  15
BB858165
LOCUS       BB858165                 502 bp    mRNA    linear   EST 26-NOV-2001
DEFINITION  BB858165 RIKEN full-length enriched, B16 F10Y cells Mus musculus
            cDNA clone G370049A15 5', mRNA sequence.
ACCESSION   BB858165
VERSION     BB858165.1  GI:17099619
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 502)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .502
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G370049A15"
                     /clone_lib="RIKEN full-length enriched, B16 F10Y cells"
                     /cell_type="B16 F10Y cells"
BASE COUNT       98 a    139 c    147 g    118 t
ORIGIN
```

```
  Query Match             13.8%;  Score 198.8;  DB 9;  Length 502;
  Best Local Similarity   68.1%;  Pred. No. 3.3e-44;
  Matches  309;  Conservative    0;  Mismatches  137;  Indels    8;  Gaps    2;

Qy    109 GCATCTCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCC 168
          | |  ||  || |||  ||||| |||  |||||||||||| |   | ||| |||||||
Db      3 GGAGGTCTGGAACCTCAGCAATGGCTACAGGGTGACTTCTTGGATCCTCCACGTTTCCAG 62

Qy    169 GCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCC 228
            ||     ||||   |  ||   | |       || |-| ||          |   || |
Db     63 ATCCTAGTGAAGACCGGTGGACCCAGC-------TGAGGACATGAGCCTATCAGCTGGGC 115

Qy    229 CGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGC 288
          |  |   |||||  ||| || || ||| | ||  ||  | ||| || || |||||  |
Db    116 TCCCTGTGCGGCCGCTGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAGGCAC 175

Qy    289 GCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTG 348
            | ||||||||| |||||    ||||||||||||||||||||| |||  ||||| |
Db    176 TGCCACCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCACAGCCAGACCTCG 235

Qy    349 GTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATC 408
          | ||  | |||||||| |||  |||||| |||||||||||||||| || |||| |||  |||
Db    236 GGCTGCCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATC 295

Qy    409 ATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGA 468
          ||||||||||||||  |||||| |  | || ||| | |       ||  |||||||||
Db    296 ATGAGAGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGC 355

Qy    469 TGTGG-CTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTC 527
          |||||  |||||  |||||||||||  ||||||||||||||  |||||||||  || ||||||
Db    356 TGTGGCCTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTC 415

Qy    528 TGGACTATTATGGAAAATCACAACCACAGCAAGG 561
          |||||  || ||||  ||  |  ||||||||  ||||
Db    416 TGGACCATCATGGGCAACTATAACCACAGTAAGG 449
```

Search completed: June 19, 2002, 07:21:40
Job time: 2586 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:35 ; Search time 352.95 Seconds
(without alignments)
7004.836 Million cell updates/sec

Title: US-09-497-957-9
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1440	100.0	1440	18	AAT96691	Hereditary haemoch
	2	1440	100.0	1440	22	AAC68429	Human hereditary h
	3	1440	100.0	2727	19	AAV23525	Haemochromatosis g
	4	1438.4	99.9	1440	22	AAC68430	Human hereditary h
	5	1438.4	99.9	1440	22	AAC68431	Human hereditary h
	6	1436.8	99.8	1440	22	AAC68432	Human hereditary h
	7	1219	84.7	2506	21	AAA96769	cDNA sequence enco
	8	559.2	38.8	596	22	AAI63897	Human polynucleoti
	9	321	22.3	10825	18	AAT96690	Hereditary haemoch
	10	321	22.3	10825	22	AAC68425	Human hereditary h
	11	321	22.3	10825	22	AAC68426	Human hereditary h
	12	321	22.3	12146	21	AAA96794	Genomic DNA of a h
c	13	321	22.3	237326	19	AAV57903	Hereditary haemoch
	14	319.4	22.2	10825	22	AAC68427	Human hereditary h
	15	319.4	22.2	10825	22	AAC68428	Human hereditary h
	16	280	19.4	359	20	AAX16055	Hereditary hemochr
	17	280	19.4	517	22	AAC68440	Human hereditary h
	18	280	19.4	5749	22	AAL36747	Human musculoskele
	19	278.4	19.3	517	22	AAC68441	Human hereditary h
	20	191.8	13.3	8622	24	ABL34142	Human immune syste
c	21	178.2	12.4	8622	24	ABL34143	Human immune syste
	22	175.2	12.2	1112	21	AAA48668	cDNA encoding chic
	23	173.6	12.1	1230	21	AAA48673	cDNA encoding chic
	24	170.4	11.8	1195	21	AAA48671	cDNA encoding chic
	25	168.4	11.7	1197	21	AAA48672	cDNA encoding chic
	26	168.4	11.7	1262	21	AAA48691	Consensus DNA of c
	27	166.8	11.6	1173	21	AAC78071	Human cancer assoc
	28	166.8	11.6	1173	22	AAF72767	Human prostate can
	29	163.6	11.4	1145	21	AAA48667	cDNA encoding chic
	30	162	11.2	1230	21	AAA48669	cDNA encoding chic
	31	162	11.2	1284	9	AAN80603	Probe F10 of Major
	32	160.4	11.1	1230	21	AAA48665	cDNA encoding chic
	33	160.4	11.1	1230	21	AAA48670	cDNA encoding chic
	34	158.8	11.0	1230	21	AAA48666	cDNA encoding chic
	35	157.2	10.9	1230	21	AAA48664	cDNA encoding chic
	36	145.2	10.1	1554	22	AAI93004	Human polynucleoti
	37	144.8	10.1	1101	12	AAQ12116	HLA-C exon Cb-1.
	38	143.6	10.0	1567	22	AAH98676	Human EST-derived
	39	143.4	10.0	1073	22	AAH42223	Nucleotide sequenc
	40	142.8	9.9	1098	22	AAS07697	Human cDNA encodin
	41	141.6	9.8	1101	12	AAQ12117	HLA-C exon Cb-2.
	42	141	9.8	2034	23	AAS90913	DNA encoding novel
	43	141	9.8	2037	23	AAS90740	DNA encoding novel
	44	140	9.7	4965	16	AAQ75973	pHLA-B7/beta-2 mic
	45	139.4	9.7	1089	22	AAH45555	Human cancer cell

ALIGNMENTS

```
RESULT   1
AAT96691
ID   AAT96691 standard; cDNA; 1440 BP.
XX
AC   AAT96691;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene cDNA clone.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             222.1268
FT                   /*tag=  a
FT   mutation        408
FT                   /*tag=  g
FT                   /note= "C to G substitution (24d2 mutation)
FT                        results in His to Asp substitution"
FT   variation       414
FT                   /*tag=  h
FT                   /note= "A to T substitution (24d7 variant)
FT                        results in Ser to Cys substitution"
FT   mutation        1066
FT                   /*tag=  i
FT                   /note= "G to A substitution (24d1 mutation
FT                        associated with HH), results in Cys to
FT                        Tyr substitution"
XX
```

PN WO9738137-A1.
XX
PD 16-OCT-1997.
XX
PF 04-APR-1997; 97WO-US06254.
XX
PR 23-MAY-1996; 96US-0652265.
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
XX
PA (MERC-) MERCATOR GENETICS INC.
XX
PI Drayna DT, Feder JN, Gnirke A, Ruddy D, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
DR WPI; 1997-512743/47.
DR P-PSDB; AAW36499.
XX
PT Hereditary haemochromatosis gene and variants - useful for diagnosis
PT and treatment of hereditary haemochromatosis disease
XX
PS Disclosure; Fig 4; 115pp; English.
XX
CC This cDNA clone, designated cDNA24, is derived from human gene
CC whose mutated form is associated with hereditary haemochromatosis
CC (HH). It was obtained from a directionally cloned plasmid-based
CC cDNA library following identification of the HH locus in the HLA
CC region of chromosome 6. A single mutation (24d1) in the HH gene
CC appears responsible for the majority of HH disease. This comprises
CC a G to A substitution that is present in 86% of affected
CC chromosomes and in 4% of unaffected chromosomes. It results in a
CC Cys to Tyr substitution in the encoded protein (see AAW36499) at a
CC critical disulphide bridge important for secondary structure. The
CC following are claimed: a 10825 bp genomic DNA sequence (I) (see
CC AAT96690), the 1437 bp cDNA sequence (Ia) and their 24d1, 24d2 and
CC 24d7 variants; a cloning or expression vector; host cells; a
CC peptide product chosen from the HH gene product, its variants
CC (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC residues of these; an antibody produced using the peptide; a method
CC to determine the presence or absence of the common HH gene
CC mutation; an animal model for the HH disease; metal chelation
CC agents, T-cell differentiation factors and therapeutic agents for
CC the mitigation of injury due to oxidative process in vivo or
CC mitigation of iron overload; a method for screening potential
CC therapeutic agents for activity in connection with HH disease; an
CC antisense oligonucleotide directed against a transcriptional
CC product of a nucleic acid sequence as above; and oligonucleotides
CC or pairs of oligonucleotides covering a range of nucleotides from
CC (I), (Ia) or their variants, useful for detecting a polymorphism in
CC the HH gene. The invention also relates to methods for screening
CC for HH homozygotes, to HH diagnosis, prenatal screening and
CC diagnosis, and therapies of HH disease, including gene therapy,
CC protein- and antibody-based therapeutics, and small molecule
CC therapeutics.
XX
SQ Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;

Query Match 100.0%; Score 1440; DB 18; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ggataaccttggctgtaccccctggggaagagcagagatatacgtgccaggtggagcacc 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 caggcctggatcagccccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260
```

```
Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440


RESULT   2
AAC68429
ID   AAC68429 standard; DNA; 1440 BP.
XX
AC   AAC68429;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis cDNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ss.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;


  Query Match             100.0%;  Score 1440;  DB 22;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ggataaccttggctgtaccccctggggaagagcagagatatacgtgccaggtggagcacc 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 caggcctggatcagccccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440


RESULT   3
AAV23525
ID   AAV23525 standard; mRNA; 2727 BP.
XX
AC   AAV23525;
XX
DT   10-JUL-1998  (first entry)
XX
DE   Haemochromatosis gene.
XX
KW   Hereditary haemochromatosis; HC gene; HH identification; diagnosis;
KW   autosomal recessive disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO9807884-A1.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-AU00539.
XX
PR   03-SEP-1996;   96AU-0002083.
PR   23-AUG-1996;   96AU-0001849.
XX
PA   (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
PI   Busfield F,  Cullen LM,  Jazwinska EC,  Powell LW;
XX
DR   WPI; 1998-179064/16.
XX
PT   Detection of autosomal recessive disorder - particularly hereditary
PT   haemochromatosis, by detecting a mutation in the HC gene
XX
PS   Disclosure; Page -; 32pp; English.
XX
CC   This sequence represents the haemochromatosis (HC) gene. Mutations in
CC   this sequence are detected using the method of the invention. The method
CC   is for identifying an individual with hereditary haemochromatosis (HH) or
CC   a predisposition to develop HH or to genetically pass on HH to an
CC   offspring, comprising isolating a biological sample and amplifying a
CC   region of genomic DNA in the biological sample encompassing all or part
CC   of the DNA between markers D6S265 and D6S276, and detecting at least one
CC   homozygous or heterozygous mutation in a nucleotide within the region.
CC   The method can also be used for identifying an individual with an
CC   autosomal recessive disorder (ARD) or predisposition to develop and/or
CC   genetically pass on an ARD to an offspring, comprises isolating a
CC   biological sample from the individual and screening genomic DNA in the
CC   sample for the presence of a homozygous or heterozygous mutation in a
CC   gene, the normal function of which, is required to prevent progression of
CC   the disorder. The method(s) can be used to identify individuals that are
CC   homozygous or heterozygous (carriers) for the mutation causing the ARD.
CC   Especially the method is used to diagnose HH or predisposition to HH by
CC   detecting a Cys282Tyr substitution. Individuals homozygous for this
CC   mutation have HH and heterozygotes are potential carriers of the
CC   disease.
XX
SQ   Sequence 2727 BP; 702 A; 606 C; 660 G; 759 T; 0 other;
```

```
  Query Match             100.0%;  Score 1440;  DB 19;  Length 2727;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020
```

```
Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtgccaggtggagcacc 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

RESULT 4
AAC68430
ID AAC68430 standard; DNA; 1440 BP.
XX
AC AAC68430;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1 mutation cDNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 348 A; 355 C; 406 G; 331 T; 0 other;

Query Match 99.9%; Score 1438.4; DB 22; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgccccccaaaag 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ccttgaactactaccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960
```

```
Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtaccaggtggagcacc 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

```
RESULT   5
AAC68431
ID   AAC68431 standard; DNA; 1440 BP.
XX
AC   AAC68431;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation cDNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ss.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 1440 BP; 347 A; 354 C; 408 G; 331 T; 0 other;
```

```
  Query Match             99.9%;  Score 1438.4;  DB 22;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccccaaaag 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtcgcc 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
```

```
Db    901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
Db    961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
Db   1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtgccaggtggagcacc 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db   1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
Db   1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

RESULT 6
AAC68432
ID AAC68432 standard; DNA; 1440 BP.
XX
AC AAC68432;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1/2 mutation cDNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 348 A; 354 C; 407 G; 331 T; 0 other;

Query Match 99.8%; Score 1436.8; DB 22; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db      1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
Db     61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
Db    121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
Db    181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
Db    241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db    301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
Db    361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtcgcc 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db    421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db    481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db    541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
Db    601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
Db    661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db    721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db    781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
```

```
Db    841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtaccaggtggagcacc 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440


RESULT   7
AAA96769
ID   AAA96769 standard; cDNA; 2506 BP.
XX
AC   AAA96769;
XX
DT   19-FEB-2001  (first entry)
XX
DE   cDNA sequence encoding a histocompatibility iron loading (HFE) protein.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1044
FT                   /*tag= a
FT                   /product= "histocompatibility iron loading (HFE) protein"
FT   sig_peptide     1..66
FT                   /*tag= b
FT   mutation        187
FT                   /*tag= c
FT                   /note= "if this base is mutated to G, then the
FT                         protein contains the mutation H63D"
FT   mutation        193
FT                   /*tag= d
FT                   /note= "if this base is mutated to T, then the
FT                         protein contains the mutation S65C"
FT   mutation        277
FT                   /*tag= e
FT                   /note= "if this base is mutated to C, then the
FT                         protein contains the mutation G93R"
FT   mutation        314
FT                   /*tag= f
FT                   /note= "if this base is mutated to C, then the
FT                         protein contains the mutation I105T, which
FT                         is associated with an iron overload disorder"
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
DR   P-PSDB; AAB19149.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Disclosure; Page 2-3; 55pp; English.
XX
CC   The present sequence encodes a human histocompatibility iron loading
CC   (HFE) protein. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 2506 BP; 648 A; 552 C; 596 G; 710 T; 0 other;


  Query Match            84.7%;  Score 1219;  DB 21;  Length 2506;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgggcccgcgagccaggccggcgcttctcctcctgatgcttttgcagaccgcggtcctg 60

Qy    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 caggggcgcttgctgcgttcacactctctgcactacctcttcatgggtgcctcagagcag 120

Qy    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gaccttggtctttccttgtttgaagctttgggctacgtggatgaccagctgttcgtgttc 180

Qy    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tatgatcatgagagtcgccgtgtggagccccgaactccatgggtttccagtagaatttca 240

Qy    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 agccagatgtggctgcagctgagtcagagtctgaaagggtgggatcacatgttcactgtt 300

Qy    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gacttctggactattatggaaaatcacaaccacagcaaggagtcccacaccctgcaggtc 360

Qy    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641
```

```
Db    361 atcctgggctgtgaaatgcaagaagacaacagtaccgagggctactggaagtacgggtat 420

Qy    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701
Db    421 gatgggcaggaccaccttgaattctgccctgacacactggattggagagcagcagaaccc 480

Qy    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761
Db    481 agggcctggcccaccaagctggagtgggaaaggcacaagattcgggccaggcagaacagg 540

Qy    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821
Db    541 gcctacctggagagggactgccctgcacagctgcagcagttgctggagctggggagaggt 600

Qy    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881
Db    601 gttttggaccaacaagtgcctcctttggtgaaggtgacacatcatgtgacctcttcagtg 660

Qy    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941
Db    661 accactctacggtgtcgggccttgaactactaccccagaacatcaccatgaagtggctg 720

Qy    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
Db    721 aaggataagcagccaatggatgccaaggagttcgaacctaaagacgtattgcccaatggg 780

Qy   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061
Db    781 gatgggacctaccagggctggataaccttggctgtaccccctggggaagagcagagatat 840

Qy   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121
Db    841 acgtgccaggtggagcacccaggcctggatcagcccctcattgtgatctgggagccctca 900

Qy   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181
Db    901 ccgtctggcaccctagtcattggagtcatcagtggaattgctgttttgtcgtcatcttg 960

Qy   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241
Db    961 ttcattggaattttgttcataatattaaggaagaggcagggttcaagaggagccatgggg 1020

Qy   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301
Db   1021 cactacgtcttagctgaacgtgagtgacacgcagcctgcagactcactgtgggaaggaga 1080

Qy   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361
Db   1081 caaaactagagactcaaagagggagtgcatttatgagctcttcatgtttcaggagagagt 1140

Qy   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421
Db   1141 tgaacctaaacatagaaattgcctgacgaactccttgattttagccttctctgttcattt 1200

Qy   1422 CCTCAAAAAGATTTCCCCA 1440
Db   1201 cctcaaaaagatttcccca 1219


RESULT   8
AAI63897
ID   AAI63897 standard; cDNA; 596 BP.
XX
AC   AAI63897;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 105.
XX
KW   Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW   neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW   antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
KW   antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW   cardiovascular disorder; neurological disease; infection; human; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155308-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01309.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
```

PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488781/53.
DR P-PSDB; AAM43591.
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders -
XX
PS Claim 1; SEQ ID NO 105; 664pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI63803-AAI64012) and
CC the encoded proteins (AAM434497-AAM43660) useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. The
CC genes were isolated from a range of human tissues disclosed in the
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 596 BP; 133 A; 157 C; 175 G; 126 T; 5 other;

Query Match 38.8%; Score 559.2; DB 22; Length 596;
Best Local Similarity 97.8%; Pred. No. 3e-151;
Matches 572; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 166 CCCGCCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGG 225
Db 13 cccgcccccaaaagaagcggagatttaacggggacgtgcggccagagctggggaaatgg 72

Qy 226 GCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGG 285
Db 73 gcccgcgagccaggccggcgcttctcctcctgatgcttttgcakaccgcggtcctgcagg 132

Qy 286 GGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACC 345
Db 133 ggcgcttgctgcgttcacactctctgcactacctcttcatgggtgcctcagagcaggacc 192

Qy 346 TTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATG 405
Db 193 ttggtctttccttgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatg 252

Qy 406 ATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCC 465
Db 253 atcatgagagtcgccgtgtggagccccgaactccatgggtttccagtagaatttcaagcc 312

Qy 466 AGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACT 525
Db 313 agatgtggctgcagctgagtcagagtctgaaa-ggtgggatcacatgttcactgttgact 371

Qy 526 TCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCC 585
Db 372 tctggactattatggraaatcacaaccacagcaaggrgtyccacaccctgcagtcatcct 431

Qy 586 TGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATG 645
Db 432 ggggctgtgaaatgcaagaagacaacagtaccgagggctactggaagtacgggtatgatg 491

```
Qy    646 GGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGG 705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 ggcaggaccaccttgaattctgccctgacacactggattggagagcagcagaacccaggg 551

Qy    706 CCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCA 750
          |||||||||||||| |||||||||||||||||||||||||||||||
Db    552 cctggcccaccaanctggagtgggaaaggcacaagattcgggcca 596


RESULT   9
AAT96690
ID   AAT96690 standard; DNA; 10825 BP.
XX
AC   AAT96690;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             361..7147
FT                   /*tag=  a
FT                   /note= "contains introns"
FT   intron          437..3761
FT                   /*tag=  b
FT                   /number= 1
FT   intron          4026..4234
FT                   /*tag=  c
FT                   /number= 2
FT   intron          4511..5605
FT                   /*tag=  d
FT                   /number= 3
FT   intron          5882..6039
FT                   /*tag=  e
FT                   /number= 4
FT   intron          6154..7106
FT                   /*tag=  f
FT                   /number= 5
FT   mutation        3872
FT                   /*tag=  g
FT                   /note= "C to G substitution (24d2 mutation)
FT                          results in His to Asp substitution"
FT   variation       3878
FT                   /*tag=  h
FT                   /note= "A to T substitution (24d7 variant)
FT                          results in Ser to Cys substitution"
FT   mutation        5834
FT                   /*tag=  i
FT                   /note= "G to A substitution (24d1 mutation
FT                          associated with HH), results in Cys to
FT                          Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   P-PSDB; AAW36499.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 3; 115pp; English.
XX
CC   This genomic DNA sequence corresponds to the human gene whose
CC   mutated form is associated with hereditary haemochromatosis (HH).
CC   To identify this novel gene, allelic association patterns were
CC   determined between known markers and the HH locus in the HLA region
CC   of chromosome 6.  A physical clone coverage was then generated
CC   extending from D6S265, which is a marker that is centromeric of
CC   HLA-A, in a telomeric direction through D6S276, a marker at which
CC   the allelic association was no longer observed.  A single mutation
CC   (24d1) in the HH gene appears responsible for the majority of HH
CC   disease.  This comprises a G to A substitution that is present in
CC   86% of affected chromosomes and in 4% of unaffected chromosomes.
CC   It results in a Cys to Tyr substitution in the encoded protein (see
CC   AAW36499) at a critical disulphide bridge important for secondary
CC   structure.  The following are claimed: the HH genomic DNA (I), a
CC   1437 bp cDNA sequence (Ia) (see AAT96691) and their 24d1, 24d2 and
CC   24d7 variants; a cloning or expression vector; host cells; a
CC   peptide product chosen from the HH gene product, its variants
CC   (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC   residues of these; an antibody produced using the peptide; a method
CC   to determine the presence or absence of the common HH gene
CC   mutation; an animal model for the HH disease; metal chelation
CC   agents, T-cell differentiation factors and therapeutic agents for
CC   the mitigation of injury due to oxidative process in vivo or
CC   mitigation of iron overload; a method for screening potential
CC   therapeutic agents for activity in connection with HH disease; an
CC   antisense oligonucleotide directed against a transcriptional
CC   product of a nucleic acid sequence as above; and oligonucleotides
CC   or pairs of oligonucleotides covering a range of nucleotides from
CC   (I), (Ia) or their variants, useful for detecting a polymorphism in
CC   the HH gene.  The invention also relates to methods for screening
CC   for HH homozygotes, to HH diagnosis, prenatal screening and
CC   diagnosis, and therapies of HH disease, including gene therapy,
CC   protein- and antibody-based therapeutics, and small molecule
CC   therapeutics.
XX
SQ   Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;


  Query Match             22.3%;  Score 321;  DB 18;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 8.6e-82;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||||
Db    4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy     561 ---------------------------------------------------------- 560

Db    4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121
```

```
Qy    561 ------------------------------------------------------------ 560

Db   4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db   4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy    629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db   4482 gctggggagaggtgttttggaccaacaag 4510
```

```
RESULT  10
AAC68425
ID   AAC68425 standard; DNA; 10825 BP.
XX
AC   AAC68425;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36869.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;
```

```
  Query Match             22.3%;  Score 321;  DB 22;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 8.6e-82;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 3881

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy    538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
          |||||||||||||||||||||||
Db   4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy    561 ------------------------------------------------------------ 560

Db   4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy    561 ------------------------------------------------------------ 560

Db   4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db   4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy    629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db   4482 gctggggagaggtgttttggaccaacaag 4510
```

```
RESULT  11
AAC68426
ID   AAC68426 standard; DNA; 10825 BP.
XX
AC   AAC68426;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1 mutation DNA.
XX
```

```
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36870.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;


  Query Match             22.3%;  Score 321;  DB 22;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 8.6e-82;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||
Db    4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy     561 ---------------------------------------------------------- 560

Db    4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy     561 ---------------------------------------------------------- 560

Db    4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                               ||||||||

Db    4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db    4482 gctggggagaggtgttttggaccaacaag 4510


RESULT  12
AAA96794
ID   AAA96794 standard; cDNA; 12146 BP.
XX
AC   AAA96794;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1028..1324
FT                   /*tag= a
FT                   /number= 1
FT   intron          1325..4651
FT                   /*tag= b
FT                   /number= 1
FT   exon            4652..4915
FT                   /*tag= c
FT                   /number= 2
FT   intron          4916..5124
FT                   /*tag= d
FT                   /number= 2
FT   exon            5125..5400
FT                   /*tag= e
FT                   /number= 3
FT   intron          5401..6493
FT                   /*tag= f
FT                   /number= 3
FT   exon            6494..6769
FT                   /*tag= g
FT                   /number= 4
FT   intron          6770..6927
FT                   /*tag= h
FT                   /number= 4
FT   exon            6928..7041
FT                   /*tag= i
FT                   /number= 5
FT   intron          7042..7994
FT                   /*tag= j
FT                   /number= 5
FT   exon            7995..9050
FT                   /*tag= k
FT                   /number= 6
```

```
FT   intron          9051..10205
FT                   /*tag=  1
FT                   /number= 6
FT   exon            10206..10637
FT                   /*tag=  m
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Example 1; Page 21-28; 55pp; English.
XX
CC   The present sequence represents the human histocompatibility iron
CC   loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;


  Query Match             22.3%;  Score 321;  DB 21;  Length 12146;
  Best Local Similarity   72.1%;  Pred. No. 9.1e-82;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4652 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 4711

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4712 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 4771

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4772 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 4831

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4832 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4891

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||||
Db    4892 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4951

Qy     561 ------------------------------------------------------------ 560

Db    4952 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 5011

Qy     561 ------------------------------------------------------------ 560

Db    5012 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 5071

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                               ||||||||
Db    5072 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 5131

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5132 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 5191

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5192 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 5251

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5252 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 5311

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5312 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 5371

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db    5372 gctggggagaggtgttttggaccaacaag 5400


RESULT  13
AAV57903/c
ID   AAV57903 standard; DNA; 237326 BP.
XX
AC   AAV57903;
XX
DT   21-DEC-1998  (first entry)
XX
DE   Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW   Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW   diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW   BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW   type 1 sodium transport gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO9814466-A1.
XX
PD   09-APR-1998.
XX
PF   30-SEP-1997;   97WO-US17658.
XX
PR   07-MAY-1997;   97US-0852495.
PR   01-OCT-1996;   96US-0724394.
XX
PA   (PROG-) PROGENTIOR INC.
XX
PI   Feder JN,  Kronmal GS,  Lauer PM,  Ruddy DA,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1998-240014/21.
XX
PT   Hereditary haemochromatosis gene products - used to develop products
PT   for the diagnosis and treatment of hereditary disorders in iron
PT   metabolism
XX
PS   Claim 1; Fig 9; 209pp; English.
XX
CC   The present invention describes hereditary haemochromatosis gene
CC   products from the human haemochromatosis gene. The present sequence
CC   represents a hereditary haemochromatosis subregion from an hereditary
CC   haemochromatosis (HH) affected individual. Also described is a
CC   method to determine the presence or absence of the common hereditary
CC   haemochromatosis (HFE) gene mutation in an individual comprising:
CC   (a) providing DNA or RNA from the individual; and (b) assessing the
```

```
CC   DNA or RNA for the presence or absence of a haplotype or genotype where
CC   the presence or absence of the haplotype genotype indicates the likely
CC   presence of the HFE gene mutation in the genome of the individual. The
CC   HFE gene sequences from the present invention can be used to develop
CC   products for use in the diagnosis and treatment of HFE. The present
CC   invention also describes BTF genes, which are homologues of the milk
CC   protein butyrophilin (BT), and can be used in the production of agonists
CC   and antagonists of BT function. Also described are: (1) a RoRet gene
CC   which can be used to develop products for the study, diagnosis and
CC   treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC   which are homologues of a type 1 sodium transport gene, and can
CC   similarly be used for hypophosphatemia.
XX
SQ   Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;


  Query Match             22.3%;  Score 321;  DB 19;  Length 237326;
  Best Local Similarity   72.1%;  Pred. No. 3.9e-81;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  43338 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 43279

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  43278 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 43219

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  43218 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 43159

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  43158 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 43099

Qy    538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
          |||||||||||||||||||||||
Db  43098 TGGAAAATCACAACCACAGCAAGGGTATGTGGAGAGGGGGCCTCACCTTCCTGAGGTTGT 43039

Qy    561 ------------------------------------------------------------ 560

Db  43038 CAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGGAAGTCTGAGGTCTTGTG 42979

Qy    561 ------------------------------------------------------------ 560

Db  42978 GGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGATGGTGG 42919

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db  42918 AAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCA 42859

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42858 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 42799

Qy    629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42798 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 42739

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42738 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 42679

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42678 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 42619

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          |||||||||||||||||||||||||||||
Db  42618 GCTGGGGAGAGGTGTTTTGGACCAACAAG 42590
```

```
RESULT  14
AAC68427
ID   AAC68427 standard; DNA; 10825 BP.
XX
AC   AAC68427;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36871.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;


  Query Match             22.2%;  Score 319.4;  DB 22;  Length 10825;
  Best Local Similarity   72.0%;  Pred. No. 2.5e-81;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels  209;  Gaps    1;

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtc 3881

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy    538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
          |||||||||||||||||||||||
Db   4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061
```

```
Qy    561 ------------------------------------------------------------ 560

Db   4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy    561 ------------------------------------------------------------ 560

Db   4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db   4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy    629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db   4482 gctggggagaggtgttttggaccaacaag 4510
```

```
RESULT  15
AAC68428
ID   AAC68428 standard; DNA; 10825 BP.
XX
AC   AAC68428;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36872.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2999 A; 2252 C; 2648 G; 2926 T; 0 other;
```

```
  Query Match             22.2%;  Score 319.4;  DB 22;  Length 10825;
  Best Local Similarity   72.0%;  Pred. No. 2.5e-81;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels  209;  Gaps    1;

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtc 3881

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy    538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||||
Db   4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy    561 ------------------------------------------------------------ 560

Db   4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy    561 ------------------------------------------------------------ 560

Db   4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db   4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy    629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db   4482 gctggggagaggtgttttggaccaacaag 4510
```

Search completed: June 19, 2002, 08:49:52
Job time: 7877 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:35 ; Search time 79.08 Seconds
(without alignments)
4472.841 Million cell updates/sec

Title: US-09-497-957-9
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1440	100.0	1440	3	US-08-652-265-9	Sequence 9, Appli
2	1440	100.0	1440	3	US-08-834-497A-9	Sequence 9, Appli
3	1440	100.0	1440	4	US-09-503-444A-9	Sequence 9, Appli
4	1438.4	99.9	1440	3	US-08-652-265-10	Sequence 10, Appl
5	1438.4	99.9	1440	3	US-08-652-265-11	Sequence 11, Appl
6	1438.4	99.9	1440	3	US-08-834-497A-10	Sequence 10, Appl
7	1438.4	99.9	1440	3	US-08-834-497A-11	Sequence 11, Appl
8	1438.4	99.9	1440	4	US-09-503-444A-10	Sequence 10, Appl
9	1438.4	99.9	1440	4	US-09-503-444A-11	Sequence 11, Appl
10	1436.8	99.8	1440	3	US-08-652-265-12	Sequence 12, Appl
11	1436.8	99.8	1440	3	US-08-834-497A-12	Sequence 12, Appl
12	1436.8	99.8	1440	4	US-09-503-444A-12	Sequence 12, Appl
13	1219	84.7	2506	4	US-09-277-457-1	Sequence 1, Appli
14	321	22.3	10825	3	US-08-652-265-1	Sequence 1, Appli
15	321	22.3	10825	3	US-08-652-265-3	Sequence 3, Appli
16	321	22.3	10825	3	US-08-834-497A-1	Sequence 1, Appli
17	321	22.3	10825	3	US-08-834-497A-3	Sequence 3, Appli
18	321	22.3	10825	4	US-09-503-444A-1	Sequence 1, Appli
19	321	22.3	10825	4	US-09-503-444A-3	Sequence 3, Appli
20	321	22.3	12146	4	US-09-277-457-27	Sequence 27, Appl
21	321	22.3	246240	2	US-08-724-394A-20	Sequence 20, Appl
22	321	22.3	246240	2	US-08-724-394A-21	Sequence 21, Appl
23	321	22.3	246240	2	US-08-724-394A-22	Sequence 22, Appl
24	319.4	22.2	10825	3	US-08-652-265-5	Sequence 5, Appli
25	319.4	22.2	10825	3	US-08-652-265-7	Sequence 7, Appli
26	319.4	22.2	10825	3	US-08-834-497A-5	Sequence 5, Appli
27	319.4	22.2	10825	3	US-08-834-497A-7	Sequence 7, Appli
28	319.4	22.2	10825	4	US-09-503-444A-5	Sequence 5, Appli
29	319.4	22.2	10825	4	US-09-503-444A-7	Sequence 7, Appli
30	280	19.4	360	3	US-08-905-124-5	Sequence 5, Appli
31	280	19.4	517	1	US-08-632-673B-3	Sequence 3, Appli
32	280	19.4	517	3	US-08-652-265-20	Sequence 20, Appl
33	280	19.4	517	3	US-08-834-497A-20	Sequence 20, Appl
34	280	19.4	517	4	US-09-503-444A-20	Sequence 20, Appl
35	279	19.4	517	1	US-08-632-673B-13	Sequence 13, Appl
36	278.4	19.3	517	1	US-08-632-673B-4	Sequence 4, Appli
37	278.4	19.3	517	3	US-08-652-265-21	Sequence 21, Appl
38	278.4	19.3	517	3	US-08-834-497A-21	Sequence 21, Appl
39	278.4	19.3	517	4	US-09-503-444A-21	Sequence 21, Appl
40	175.2	12.2	1112	3	US-08-890-719-5	Sequence 5, Appli
41	173.6	12.1	1230	3	US-08-890-719-9	Sequence 9, Appli
42	170.4	11.8	1195	3	US-08-890-719-10	Sequence 10, Appl
43	168.4	11.7	1197	3	US-08-890-719-7	Sequence 7, Appli
44	168.4	11.7	1262	3	US-08-890-719-37	Sequence 37, Appl
45	163.6	11.4	1145	3	US-08-890-719-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-08-652-265-9
; Sequence 9, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
;   APPLICANT: Thomas, Winston J.
;   APPLICANT: Drayna, Dennis T.
;   APPLICANT: Feder, John N.
;   APPLICANT: Gnirke, Andreas
;   APPLICANT: Ruddy, David
;   APPLICANT: Tsuchihashi, Zenta
;   APPLICANT: Wolff, Roger K.
;   TITLE OF INVENTION: Hereditary Hemochromatosis Gene
;   NUMBER OF SEQUENCES: 44
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, Eighth Floor
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/652,265
;     FILING DATE: 23-MAY-1996
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Smith, William M.
;     REGISTRATION NUMBER: 30,223
;     REFERENCE/DOCKET NUMBER: 17957-000500
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 576-0200
;     TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 9:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1440 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 222..1268
;   FEATURE:
;     NAME/KEY: allele
;     LOCATION: replace(408, "c")
```

```
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(414, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(1066, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d1
US-08-652-265-9

  Query Match             100.0%;  Score 1440;  DB 3;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440

RESULT   2
US-08-834-497A-9
; Sequence 9, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-9


  Query Match             100.0%;  Score 1440;  DB 3;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
```

```
Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   3
US-09-503-444A-9
; Sequence 9, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-9


  Query Match             100.0%;  Score 1440;  DB 4;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
```

```
Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT   4
US-08-652-265-10
; Sequence 10, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-10


  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
```

```
Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   5
US-08-652-265-11
; Sequence 11, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-652-265-11


  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   6
US-08-834-497A-10
; Sequence 10, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-10


  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
```

```
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT   7
US-08-834-497A-11
; Sequence 11, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
```

```
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-834-497A-11


  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   8
US-09-503-444A-10
; Sequence 10, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-10


  Query Match             99.9%;  Score 1438.4;  DB 4;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   9
US-09-503-444A-11
; Sequence 11, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
```

```
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-09-503-444A-11


  Query Match             99.9%;  Score 1438.4;  DB 4;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
```

```
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440

Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT 10
US-08-652-265-12
; Sequence 12, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-12


  Query Match           99.8%;  Score 1436.8;  DB 3;  Length 1440;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1438;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
```

```
Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

RESULT 11
US-08-834-497A-12
; Sequence 12, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(408, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1066, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
US-08-834-497A-12

Query Match 99.8%; Score 1436.8; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT  12
US-09-503-444A-12
; Sequence 12, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;  .   ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-12


  Query Match             99.8%;  Score 1436.8;  DB 4;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1438;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
```

```
Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
```

```
Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

RESULT 13
US-09-277-457-1
; Sequence 1, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
; APPLICANT: Rothenberg, Barry E.
; APPLICANT: Sawada-Hirai, Ritsuko
; APPLICANT: Barton, James C.
; TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
; FILE REFERENCE: 10653/002001
; CURRENT APPLICATION NUMBER: US/09/277,457
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2506
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: Missense mutation at nucleotide 314
US-09-277-457-1

Query Match 84.7%; Score 1219; DB 4; Length 2506;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgggcccgcgagccaggccggcgcttctcctcctgatgcttttgcagaccgcggtcctg 60

Qy    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 caggggcgcttgctgcgttcacactctctgcactacctcttcatgggtgcctcagagcag 120

Qy    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gaccttggtctttccttgtttgaagctttgggctacgtggatgaccagctgttcgtgttc 180

Qy    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tatgatcatgagagtcgccgtgtggagccccgaactccatgggtttccagtagaatttca 240

Qy    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 agccagatgtggctgcagctgagtcagagtctgaaagggtgggatcacatgttcactgtt 300

Qy    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gacttctggactattatggaaaatcacaaccacagcaaggagtcccacaccctgcaggtc 360

Qy    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 atcctgggctgtgaaatgcaagaagacaacagtaccgagggctactggaagtacgggtat 420

Qy    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gatgggcaggaccaccttgaattctgccctgacacactggattggagagcagcagaaccc 480

Qy    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 agggcctggcccaccaagctggagtgggaaaggcacaagattcgggccaggcagaacagg 540
```

```
Qy    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gcctacctggagagggactgccctgcacagctgcagcagttgctggagctggggagaggt 600

Qy    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 gttttggaccaacaagtgcctcctttggtgaaggtgacacatcatgtgacctcttcagtg 660

Qy    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 accactctacggtgtcgggccttgaactactaccccagaacatcaccatgaagtggctg 720

Qy    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 aaggataagcagccaatggatgccaaggagttcgaacctaaagacgtattgcccaatggg 780

Qy   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gatgggacctaccagggctggataaccttggctgtacccccctggggaagagcagagatat 840

Qy   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 acgtgccaggtggagcacccaggcctggatcagccccctcattgtgatctgggagccctca 900

Qy   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ccgtctggcaccctagtcattggagtcatcagtggaattgctgttttgtcgtcatcttg 960

Qy   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ttcattggaattttgttcataatattaaggaagaggcagggttcaagaggagccatgggg 1020

Qy   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 cactacgtcttagctgaacgtgagtgacacgcagcctgcagactcactgtgggaaggaga 1080

Qy   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 caaaactagagactcaaagagggagtgcatttatgagctcttcatgtttcaggagagagt 1140

Qy   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 tgaacctaaacatagaaattgcctgacgaactccttgattttagccttctctgttcattt 1200

Qy   1422 CCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||
Db   1201 cctcaaaaagatttcccca 1219


RESULT  14
US-08-652-265-1
; Sequence 1, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /note= "No. 6025130mal or wild-type (unaffected)
;      OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;      OTHER INFORMATION:  allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) allele
;      OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;      OTHER INFORMATION:  allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;      OTHER INFORMATION:  allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3878, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-1


  Query Match             22.3%;  Score 321;  DB 3;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 8.9e-83;
```

```
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3822 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------------- 560
           |||||||||||||||||||||||
Db    4002 TGGAAAATCACAACCACAGCAAGGGTATGTGGAGAGGGGGCCTCACCTTCCTGAGGTTGT 4061

Qy     561 ---------------------------------------------------------- 560

Db    4062 CAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGGAAGTCTGAGGTCTTGTG 4121

Qy     561 ---------------------------------------------------------- 560

Db    4122 GGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGATGGTGG 4181

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db    4182 AAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCA 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4302 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 4361

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4362 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 4421

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4422 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 4481

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db    4482 GCTGGGGAGAGGTGTTTTGGACCAACAAG 4510


RESULT  15
US-08-652-265-3
; Sequence 3, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-3


  Query Match             22.3%;  Score 321;  DB 3;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 8.9e-83;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3822 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 3881
```

```
Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||
Db    4002 TGGAAAATCACAACCACAGCAAGGGTATGTGGAGAGGGGGCCTCACCTTCCTGAGGTTGT 4061

Qy     561 ------------------------------------------------------------ 560

Db    4062 CAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGGAAGTCTGAGGTCTTGTG 4121

Qy     561 ------------------------------------------------------------ 560

Db    4122 GGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGATGGTGG 4181

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db    4182 AAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCA 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4302 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 4361

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4362 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 4421

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4422 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 4481

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db    4482 GCTGGGGAGAGGTGTTTTGGACCAACAAG 4510
```

Search completed: June 19, 2002, 08:40:03
Job time: 7288 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:23:48 ; Search time 3675.59 Seconds
(without alignments)
8198.470 Million cell updates/sec

Title: US-09-497-957-12
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1440	100.0	1440	6	AR117796	AR117796 Sequence
2	1440	100.0	1440	6	AR149466	AR149466 Sequence
3	1438.4	99.9	1440	6	AR117794	AR117794 Sequence
4	1438.4	99.9	1440	6	AR117795	AR117795 Sequence
5	1438.4	99.9	1440	6	AR149464	AR149464 Sequence
6	1438.4	99.9	1440	6	AR149465	AR149465 Sequence
7	1436.8	99.8	1440	6	AR117793	AR117793 Sequence
8	1436.8	99.8	1440	6	AR149463	AR149463 Sequence
9	1436.8	99.8	2727	9	HSU60319	U60319 Homo sapien
10	1197.8	83.2	1280	9	HSA249335	AJ249335 Homo sapi
11	1185.8	82.3	1200	9	AF115265	AF115265 Homo sapi
12	989.8	68.7	1045	9	AF079407	AF079407 Homo sapi
13	893.8	62.1	1320	4	AY007541	AY007541 Ceratothe
14	893.8	62.1	1320	4	AY007543	AY007543 Dicerorhi
15	893.6	62.1	1319	4	AY007544	AY007544 Rhinocero
16	890.6	61.8	2332	4	AY007542	AY007542 Diceros b
17	847.8	58.9	1085	9	HSA249336	AJ249336 Homo sapi
18	783.8	54.4	1073	9	HSA249337	AJ249337 Homo sapi
19	772	53.6	1885	9	AF144242	AF144242 Homo sapi
20	709.8	49.3	823	9	AF079408	AF079408 Homo sapi
21	634.4	44.1	1706	10	RNHERHAEM	AJ001517 Rattus no
22	619.6	43.0	1529	10	MMU66849	U66849 Mus musculu
23	615.8	42.8	781	9	AF079409	AF079409 Homo sapi
24	571.4	39.7	809	9	HSA250635	AJ250635 Homo sapi
25	546.6	38.0	987	9	AF150664	AF150664 Homo sapi
26	486.8	33.8	686	9	AF144240	AF144240 Homo sapi
27	475	33.0	789	10	AF008587	AF008587 Rattus no
28	461.6	32.1	804	9	AF149804	AF149804 Homo sapi
29	447.2	31.1	819	10	AF176534	AF176534 Rattus no
30	434.8	30.2	735	9	AF115264	AF115264 Homo sapi
31	391.4	27.2	560	9	AF144243	AF144243 Homo sapi
32	321	22.3	10825	6	AR117791	AR117791 Sequence
33	321	22.3	10825	6	AR117792	AR117792 Sequence
34	321	22.3	10825	6	AR149461	AR149461 Sequence
35	321	22.3	10825	6	AR149462	AR149462 Sequence
36	320.4	22.2	492	9	AF144241	AF144241 Homo sapi
37	320.4	22.2	874	9	HSHLAH2	Y09800 H.sapiens H
38	319.4	22.2	10825	6	AR117789	AR117789 Sequence
39	319.4	22.2	10825	6	AR117790	AR117790 Sequence
40	319.4	22.2	10825	6	AR149459	AR149459 Sequence
41	319.4	22.2	10825	6	AR149460	AR149460 Sequence
42	319.4	22.2	12146	9	HSHFE	Z92910 Homo sapien
c 43	319.4	22.2	193752	2	AL359892	AL359892 Homo sapi
44	319.4	22.2	246240	6	AR036572	AR036572 Sequence
45	319.4	22.2	246240	6	AR036573	AR036573 Sequence

ALIGNMENTS

RESULT 1
AR117796
LOCUS AR117796 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 12 from patent US 6140305.
ACCESSION AR117796
VERSION AR117796.1 GI:14098702
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 12 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. .1440
/organism="unknown"
BASE COUNT 348 a 354 c 407 g 331 t
ORIGIN

Query Match 100.0%; Score 1440; DB 6; Length 1440;

```
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   2
AR149466
LOCUS       AR149466                1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 12 from patent US 6228594.
ACCESSION   AR149466
VERSION     AR149466.1  GI:15114057
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 12 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    354 c    407 g    331 t
ORIGIN


  Query Match           100.0%;  Score 1440;  DB 6;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
```

```
Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
```

```
Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   3
AR117794
LOCUS       AR117794                1440 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 10 from patent US 6140305.
ACCESSION   AR117794
VERSION     AR117794.1  GI:14098700
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 10 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    355 c    406 g    331 t
ORIGIN


  Query Match             99.9%;  Score 1438.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   4
AR117795
LOCUS       AR117795               1440 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 11 from patent US 6140305.
ACCESSION   AR117795
VERSION     AR117795.1  GI:14098701
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 11 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"

BASE COUNT      347 a    354 c    408 g    331 t
ORIGIN


  Query Match             99.9%;  Score 1438.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
```

```
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   5
AR149464
LOCUS       AR149464               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 10 from patent US 6228594.
ACCESSION   AR149464
VERSION     AR149464.1  GI:15114055
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 10 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    355 c    406 g    331 t
ORIGIN


  Query Match             99.9%;  Score 1438.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
```

```
Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   6
AR149465
LOCUS       AR149465               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 11 from patent US 6228594.
ACCESSION   AR149465
VERSION     AR149465.1  GI:15114056
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 11 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    354 c    408 g    331 t
ORIGIN


  Query Match             99.9%;  Score 1438.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   7
AR117793
LOCUS       AR117793               1440 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 9 from patent US 6140305.
ACCESSION   AR117793
VERSION     AR117793.1  GI:14098699
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
```

```
          Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 9 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    355 c    407 g    331 t
ORIGIN
```

Query Match 99.8%; Score 1436.8; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT   8
AR149463
LOCUS       AR149463               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 9 from patent US 6228594.
ACCESSION   AR149463
VERSION     AR149463.1  GI:15114054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 9 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    355 c    407 g    331 t
ORIGIN
```

Query Match 99.8%; Score 1436.8; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
```

```
Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

RESULT 9
HSU60319
LOCUS HSU60319 2727 bp mRNA linear PRI 29-OCT-1997
DEFINITION Homo sapiens haemochromatosis protein (HLA-H) mRNA, complete cds.
ACCESSION U60319
VERSION U60319.1 GI:1469789
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2727)
AUTHORS Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A., Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A., Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P., Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C., Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E., Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J., Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
TITLE A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis
JOURNAL Nature Genet. 13 (4), 399-408 (1996)
MEDLINE 96331279
REFERENCE 2 (bases 1 to 2727)
AUTHORS Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A., Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A., Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P., Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C., Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E., Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J., Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1996) Mercator Genetics, 4040 Campbell Ave., Menlo Park, CA 94025, USA
FEATURES Location/Qualifiers
source 1. .2727
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
gene 1. .2727
/gene="HLA-H"
/note="HFE"
CDS 222. .1268
/gene="HLA-H"
/codon_start=1
/product="haemochromatosis protein"
/protein_id="AAC51823.1"
/db_xref="GI:1469790"
/translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
SRGAMGHYVLAERE"
BASE COUNT 702 a 606 c 660 g 759 t

ORIGIN

Query Match 99.8%; Score 1436.8; DB 9; Length 2727;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

RESULT 10
HSA249335
LOCUS HSA249335 1280 bp mRNA linear PRI 04-SEP-2001
DEFINITION Homo sapiens mRNA for hemochromatosis protein (HFE gene) splice variant 1.
ACCESSION AJ249335
VERSION AJ249335.1 GI:15485418
KEYWORDS alternative splicing; hemochromatosis protein; HFE gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1280)
AUTHORS OIiva,R. and Sanchez,M.
TITLE Identification of different alternative splicing forms of the HFE gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1280)
AUTHORS Oliva,R.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1999) Oliva R., Faculty of Medicine and Clinic Hospital, Human Genome Research Group, Casanova 143, 08036, SPAIN
FEATURES Location/Qualifiers
source 1. .1280
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p22"
/cell_line="HepG2"
gene 161. .1138
/gene="HFE"
CDS 161. .1138
/gene="HFE"
/function="iron metabolism"
/note="alternative splicing form with deletion of first 69 pb of exon 2"
/codon_start=1
/evidence=experimental
/product="hemochromatosis protein"
/protein_id="CAC67792.1"
/db_xref="GI:15485419"
/translation="MGPRARPALLLLMLLQTAVLQGRLLPLGYVDDQLFVFYDHESRR

```
                     VEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCE
                     MQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYL
                     ERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLK
                     DKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWEP
                     SPSGTLVIGVISGIAVFVVILFIGILFIILRKRQGSRGAMGHYVLAERE"
BASE COUNT      311 a    314 c    371 g    284 t
ORIGIN


  Query Match           83.2%;  Score 1197.8;  DB 9;  Length 1280;
  Best Local Similarity   94.7%;  Pred. No. 0;
  Matches 1278;  Conservative    0;  Mismatches    2;  Indels   69;  Gaps    1;

Qy      62 CTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAGC 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAGC 60

Qy     122 CTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGA 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGA 120

Qy     182 AGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCC 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCC 180

Qy     242 GGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTC 301
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGC---- 236

Qy     302 ACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTT 361

Db     237 ------------------------------------------------------------ 236

Qy     362 TGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCG 421
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     237 -----CTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCG 291

Qy     422 TGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCT 481
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 TGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCT 351

Qy     482 GAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGA 541
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     352 GAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGA 411

Qy     542 AAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCA 601
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     412 AAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCA 471

Qy     602 AGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGA 661
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     472 AGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGA 531

Qy     662 ATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCT 721
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     532 ATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCT 591

Qy     722 GGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTG 781
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     592 GGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTG 651

Qy     782 CCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCC 841
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     652 CCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCC 711

Qy     842 TCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGC 901
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     712 TCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGC 771

Qy     902 CTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGA 961
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     772 CTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGA 831

Qy     962 TGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTG 1021
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     832 TGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTG 891

Qy    1022 GATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACCC 1081
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     892 GATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCC 951

Qy    1082 AGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCAT 1141
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     952 AGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCAT 1011

Qy    1142 TGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCAT 1201
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1012 TGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCAT 1071

Qy    1202 AATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACG 1261
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1072 AATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACG 1131

Qy    1262 TGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA 1321
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1132 TGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA 1191

Qy    1322 GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATT 1381
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1192 GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATT 1251

Qy    1382 GCCTGACGAACTCCTTGATTTTAGCCTTC 1410
           |||||||||||||||||||||||||||||
Db    1252 GCCTGACGAACTCCTTGATTTTAGCCTTC 1280


RESULT  11
AF115265
LOCUS       AF115265                1200 bp    mRNA    linear   PRI 07-MAY-2001
DEFINITION  Homo sapiens hemochromatosis termination variant terE6 (HFE) mRNA,
            complete cds.
ACCESSION   AF115265
VERSION     AF115265.1  GI:11094314
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
  AUTHORS   Thenie,A., Orhant,M., Gicquel,I., Fergelot,P., Le Gall,J.Y.,
            David,V. and Mosser,J.
  TITLE     The HFE gene undergoes alternate splicing processes
  JOURNAL   Blood Cells Mol. Dis. 26 (2), 155-162 (2000)
  MEDLINE   20448010
   PUBMED   11001625
REFERENCE   2  (bases 1 to 1200)
  AUTHORS   Thenie,A., Orhant,M. and Mosser,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-1998) UPR 41 CNRS, Faculte de Medecine, 2, av du
            Pr. Bernard, Rennes 35043, France
FEATURES             Location/Qualifiers
     source          1. .1200
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22"
     gene            1. .1200
                     /gene="HFE"
     CDS             12. .1058
                     /gene="HFE"
                     /codon_start=1
                     /product="hemochromatosis termination variant terE6"
                     /protein_id="AAG29572.1"
                     /db_xref="GI:11094315"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
```

```
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                     SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                     EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                     SRGAMGHYVLAERE"
BASE COUNT      298 a    290 c    346 g    266 t
ORIGIN


  Query Match            82.3%;  Score 1185.8;  DB 9;  Length 1200;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1187;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    211 GAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGA 270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGA 60

Qy    271 CCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTG 330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTG 120

Qy    331 CCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGC 390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGC 180

Qy    391 TGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCA 450
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCA 240

Qy    451 GTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACA 510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACA 300

Qy    511 TGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACA 570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACA 360

Qy    571 CCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGA 630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGA 420

Qy    631 AGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAG 690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAG 480

Qy    691 CAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCA 750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCA 540

Qy    751 GGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGC 810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGC 600

Qy    811 TGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGA 870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGA 660

Qy    871 CCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCA 930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCA 720

Qy    931 TGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTAT 990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTAT 780

Qy    991 TGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAG 1050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAG 840

Qy   1051 AGCAGAGATATACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCT 1110
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCT 900

Qy   1111 GGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTG 1170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTG 960

Qy   1171 TCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAG 1230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAG 1020

Qy   1231 GAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTG 1290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTG 1080

Qy   1291 TGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTT 1350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTT 1140

Qy   1351 CAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGA 1399
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGA 1189


RESULT  12
AF079407
LOCUS       AF079407                1045 bp    mRNA    linear   PRI 18-MAR-1999
DEFINITION  Homo sapiens hemochromatosis splice variant del14E4 (HFE) mRNA,
            complete cds.
ACCESSION   AF079407
VERSION     AF079407.1  GI:3695106
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1045)
  AUTHORS   Rhodes,D.A. and Trowsdale,J.
  TITLE     Alternate splice variants of the hemochromatosis gene Hfe
  JOURNAL   Immunogenetics 49 (4), 357-359 (1999)
  MEDLINE   99180629
REFERENCE   2  (bases 1 to 1045)
  AUTHORS   Rhodes,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-1998) Immunology, University of Cambridge, Tennis
            Court Road, Cambridge CB2 1QP, UK
FEATURES             Location/Qualifiers
     source          1. .1045
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1"
     gene            1. .1045
                     /gene="HFE"
     CDS             37. .1041
                     /gene="HFE"
                     /codon_start=1
                     /product="hemochromatosis splice variant del14E4"
                     /protein_id="AAC62646.1"
                     /db_xref="GI:3695107"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVTTLRCRALNYYP
                     QNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLD
                     QPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQGSRGAMGHYVLAERE
                     "
BASE COUNT      243 a    259 c    314 g    229 t
ORIGIN


  Query Match            68.7%;  Score 989.8;  DB 9;  Length 1045;
  Best Local Similarity   96.0%;  Pred. No. 7.7e-281;
  Matches 1043;  Conservative    0;  Mismatches    2;  Indels   42;  Gaps    1;

Qy    186 GAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCG 245
```

```
Db      1 GAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCG 60

Qy    246 CTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACAC 305
Db     61 CTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACAC 120

Qy    306 TCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAA 365
Db    121 TCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAA 180

Qy    366 GCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTG 425
Db    181 GCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTG 240

Qy    426 GAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGT 485
Db    241 GAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGT 300

Qy    486 CAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAAT 545
Db    301 CAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAAT 360

Qy    546 CACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAA 605
Db    361 CACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAA 420

Qy    606 GACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTC 665
Db    421 GACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTC 480

Qy    666 TGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAG 725
Db    481 TGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAG 540

Qy    726 TGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCT 785
Db    541 TGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCT 600

Qy    786 GCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCT 845
Db    601 GCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACA---------- 650

Qy    846 TTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTG 905
Db    651 ------------------------------AGTGACCACTCTACGGTGTCGGGCCTTG 678

Qy    906 AACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCC 965
Db    679 AACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCC 738

Qy    966 AAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATA 1025
Db    739 AAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATA 798

Qy   1026 ACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACCCAGGC 1085
Db    799 ACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGC 858

Qy   1086 CTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGA 1145
Db    859 CTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGA 918

Qy   1146 GTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATA 1205
Db    919 GTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATA 978

Qy   1206 TTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAG 1265
Db    979 TTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAG 1038

Qy   1266 TGACACG 1272
Db   1039 TGACACG 1045
```

```
RESULT  13
AY007541
LOCUS       AY007541               1320 bp    mRNA    linear   MAM 22-OCT-2000
DEFINITION  Ceratotherium simum HFE protein mRNA, complete cds.
ACCESSION   AY007541
VERSION     AY007541.1  GI:10945687
KEYWORDS    .
SOURCE      white rhinoceros.
  ORGANISM  Ceratotherium simum
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
REFERENCE   1  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Ceratotherium simum"
                     /db_xref="taxon:9807"
     CDS             86. .1132
                     /codon_start=1
                     /product="HFE protein"
                     /protein_id="AAG23701.1"
                     /db_xref="GI:10945688"
                     /translation="MGPRARPALFFLILLRTVAAQGRPPRSHSLRYLFMGASERDHGL
                     PLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWLQLSQSLKGWDHMFIVDF
                     WTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDHLEFCPETLDWRAAES
                     RALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQVPPLVKVTHHVAS
                     AVTTLRCQALNFYPQNITMRWLKDRKPMDVKDAESKDVLPSGDGTYQSWEALAVPPGE
                     EQRYTCQVEHPGLDQPLTATWEPSLSNTLVTGVISGIAVCVIIFFIGILFRILRKRQA
                     SRGAMGDYVLGECE"
BASE COUNT      284 a    363 c    394 g    279 t
ORIGIN
```

```
  Query Match             62.1%;  Score 893.8;  DB 4;  Length 1320;
  Best Local Similarity   82.2%;  Pred. No. 1.8e-252;
  Matches 1063;  Conservative    0;  Mismatches  227;  Indels    3;  Gaps    3;

Qy    137 AGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACG 196
Db      2 AGAGTGACTTCTTGGGTCACTCCTATTTCCCCGCCCCACCGAAGCAGCTGA-CGCGGGCC 60

Qy    197 GGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCT 256
Db     61 AGCACCCGCGGTCCGAGCCGGGGAAATGGGCCCGCGAGCCCGACCCGCGCTCTTCTTCCT 120

Qy    257 GATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTA 316
Db    121 GATCCTCCTGCGGACCGTGGCCGCGCAGGGGCGACCACCGCGGTCACACTCTCTGCGCTA 180

Qy    317 CCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTA 376
Db    181 CCTCTTCATGGGCGCCTCAGAGCGGGACCACGGCCTGCCCCTGTTTGAAGCCTTGGGCTA 240

Qy    377 CGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGAGCCCCGAAC 436
Db    241 TGTGGATGACGAGCTTTTCGTGGCCTACAATCACGAGAGTCGCCGTGCAGAGTCCCGTGC 300

Qy    437 TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA 496
Db    301 CCAGTGGGTCTTGGGTGAGGCCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAA 360

Qy    497 AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAG 556
```

```
Db    361 AGGGTGGGATCACATGTTCATCGTCGACTTCTGGACCATCATGGACAACCACAACCACAG 420

Qy    557 CAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTAC 616
Db    421 CAAGGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGGTGCAAGAGGACAACAGTAC 480

Qy    617 CGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACAC 676
Db    481 CAGGGGTTTCTGGAAGTACGGGTACGATGGGCAGGACCATCTTGAATTCTGCCCTGAGAC 540

Qy    677 ACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCA 736
Db    541 ACTGGATTGGAGAGCAGCAGAGTCCAGGGCCCTGACCACCAAGCTGGAGTGGGAAGTGAA 600

Qy    737 CAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCA 796
Db    601 CAAGATTCGGGCCAAGCAGAACAGAGCCTATCTCGAGAGGGACTGCCCTGAGCAGCTGCA 660

Qy    797 GCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGT 856
Db    661 GTGGTTGCTGGAGCTGGGGAGAGGGGTTCTGGACCAGCAAGTGCCTCCTTTGGTGAAGGT 720

Qy    857 GACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCC 916
Db    721 GACTCATCATGTGGCCTCTGCAGTGACCACTCTACGGTGTCAGGCTCTGAACTTCTACCC 780

Qy    917 CCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGA 976
Db    781 CCAGAACATCACCATGAGGTGGTTGAAGGACAGGAAGCCAATGGATGTCAAGGACGCCGA 840

Qy    977 ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGT 1036
Db    841 GTCTAAGGATGTGCTGCCCAGCGGGGACGGGACCTACCAGAGCTGGGAGGCTTTGGCCGT 900

Qy   1037 ACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCC 1096
Db    901 GCCCCCTGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 960

Qy   1097 CCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGG 1156
Db    961 ACTCACTGCCACCTGGGAGCCCTCACTTTCTAACACCCTCGTCACTGGAGTCATCAGTGG 1020

Qy   1157 AATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAG 1216
Db   1021 GATTGCTGTTTGTGTCATCATCTTCTTCATTGGAATTTTGTTCAGAATCTTAAGGAAAAG 1080

Qy   1217 GCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGC 1276
Db   1081 ACAAGCTTCAAGAGGAGCCATGGGGGACTACGTCTTAGGTGAATGTGAGTGACGCACAGC 1140

Qy   1277 CTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATG 1336
Db   1141 CTGCAGACTCCTTGTGGGGAGGACACATAATTGGAGACTCAGAGAGGGTGCAC-TTGCGG 1199

Qy   1337 AGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCC 1395
Db   1200 CATTCTTTGTGTTTCAGGACAGAGTTGAACCTAACAAATGGGAACTGTCCCAGGAGGTCT 1259

Qy   1396 TTGATTTTAGCCTTCTCTGTTCATTTCCTCAAA 1428
Db   1260 CTGCTCTTAGCCTTCCCTGTTCACTCTCAAAAA 1292


RESULT  14
AY007543
LOCUS       AY007543               1320 bp    mRNA    linear   MAM 22-OCT-2000
DEFINITION  Dicerorhinus sumatrensis HFE protein mRNA, complete cds.
ACCESSION   AY007543
VERSION     AY007543.1  GI:10945691
KEYWORDS    .
SOURCE      Sumatran rhinoceros.
  ORGANISM  Dicerorhinus sumatrensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.
REFERENCE   1  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1320)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Dicerorhinus sumatrensis"
                     /db_xref="taxon:89632"
     CDS             86. .1132
                     /codon_start=1
                     /product="HFE protein"
                     /protein_id="AAG23703.1"
                     /db_xref="GI:10945692"
                     /translation="MGPRARPALFFLILLRTVAAQGRPPRSHSLRYLFMGASERDHGL
                     PLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWLQLSQSLKGWDHMFIVDF
                     WTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDHLEFCPETLDWRAAES
                     RALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQVPPLVKVTHHVAS
                     AVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQSWVALAVPPGE
                     EQRYTCQVEHPGLDQPLTATWEPSLSNTLVTGVISGIAVCVIIFLIGILFRILRKRQA
                     SRGAMGDYVLAECE"
BASE COUNT      281 a    358 c    398 g    283 t
ORIGIN


  Query Match           62.1%;  Score 893.8;  DB 4;  Length 1320;
  Best Local Similarity  82.2%;  Pred. No. 1.8e-252;
  Matches 1063;  Conservative    0;  Mismatches  227;  Indels    3;  Gaps    3;

Qy    137 AGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACG 196
Db      2 AGGGTGACTTCTTGGGTCACTCCTATTTTCCCGCCCCACCGAAGCAGCTGACA-CGGGCC 60

Qy    197 GGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCT 256
Db     61 AGCACCCGCGGTCCGAGCCGGGGAAATGGGCCCGCGAGCCCGACCCGCGCTGTTCTTCCT 120

Qy    257 GATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTA 316
Db    121 GATCCTCCTGCGAACCGTGGCCGCGCAGGGGCGACCGCCGCGGTCACACTCTCTGCGCTA 180

Qy    317 CCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTA 376
Db    181 CCTCTTCATGGGCGCCTCAGAGCGGGACCACGGCCTGCCCCTGTTTGAAGCCTTGGGCTA 240

Qy    377 CGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGAGCCCCGAAC 436
Db    241 TGTGGATGACGAGCTTTTCGTGGCGTACAATCATGAGAGTCGCCGTGCAGAGTCCCGTGC 300

Qy    437 TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA 496
Db    301 CCAGTGGGTCTTGGGTGAGGCCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAA 360

Qy    497 AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAG 556
Db    361 AGGGTGGGATCACATGTTCATCGTCGACTTCTGGACCATCATGGACAACCACAACCACAG 420

Qy    557 CAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTAC 616
Db    421 CAAGGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGGTGCAAGAGGACAACAGCAC 480

Qy    617 CGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACAC 676
Db    481 CAGGGGGTTCTGGAAGTACGGGTACGATGGGCAGGACCATCTTGAATTCTGCCCTGAGAC 540

Qy    677 ACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCA 736
```

```
Db    541 ACTGGATTGGAGAGCAGCAGAGTCCAGGGCGCTGACCACCAAGCTGGAGTGGGAAGTGAA 600

Qy    737 CAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCA 796
Db    601 CAAGATTCGGGCCAAGCAGAACAGAGCCTATCTCGAGAGAGACTGCCCTGAGCAGCTGCA 660

Qy    797 GCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGT 856
Db    661 GTGGTTGCTGGAGCTGGGGAGAGGGGTTCTGGACCAGCAAGTGCCTCCTTTGGTGAAGGT 720

Qy    857 GACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCC 916
Db    721 GACTCATCATGTGGCCTCTGCAGTGACCACTCTACGGTGTCAGGCTCTGAACTTCTACCC 780

Qy    917 CCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGA 976
Db    781 CCAGAACATCACCATGAGGTGGTTGAAGGACAGGAAGCCAGTGGATGTCAAGGACGCCGA 840

Qy    977 ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGT 1036
Db    841 GTCTAAGGATGTGCTGCCCAGCGGGGATGGGACCTACCAGAGCTGGGTGGCTTTGGCCGT 900

Qy   1037 ACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCC 1096
Db    901 GCCCCCTGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 960

Qy   1097 CCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGG 1156
Db    961 CCTCACTGCCACCTGGGAGCCCTCACTTTCTAACACCCTTGTCACTGGAGTCATCAGTGG 1020

Qy   1157 AATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAG 1216
Db   1021 GATTGCTGTTTGTGTCATCATCTTCCTCATTGGAATTTTGTTCAGAATCTTAAGGAAAAG 1080

Qy   1217 GCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGC 1276
Db   1081 ACAAGCTTCGAGAGGAGCCATGGGGGACTACGTCTTAGCTGAATGTGAGTGAAGCACAGC 1140

Qy   1277 CTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATG 1336
Db   1141 CTGCAGACTCCTTGTGGGGAGGACACATAATTGGAGACTCAGAGAGGGTGCAC-TTGTGG 1199

Qy   1337 AGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCC 1395
Db   1200 CGTTCTTCGTGTTTCAGGACAGAGTTGAACCTAACAAATGGGAACTGTCCCAGGAGGTCT 1259

Qy   1396 TTGATTTTAGCCTTCTCTGTTCATTTCCTCAAA 1428
Db   1260 CTGCTCTTAGCCTTCCCTGTTCACTCTCAAAAA 1292


RESULT  15
AY007544
LOCUS       AY007544               1319 bp    mRNA    linear   MAM 22-OCT-2000
DEFINITION  Rhinoceros unicornis HFE protein mRNA, complete cds.
ACCESSION   AY007544
VERSION     AY007544.1  GI:10945693
KEYWORDS    .
SOURCE      greater Indian rhinoceros.
  ORGANISM  Rhinoceros unicornis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
REFERENCE   1  (bases 1 to 1319)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1319)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA

FEATURES             Location/Qualifiers
     source          1. .1319
                     /organism="Rhinoceros unicornis"
                     /db_xref="taxon:9809"
     CDS             86. .1132
                     /codon_start=1
                     /product="HFE protein"
                     /protein_id="AAG23704.1"
                     /db_xref="GI:10945694"
                     /translation="MGPRARPALPFLILLRTVAAQGRPPRSHSLRYLFMGASERDHGL
                     PLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWLQLSQSLKGWDHMFIVDF
                     WTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWMYGYDGQDHLEFCPETLDWRAAES
                     RALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQVPPLVKVTHHVAS
                     AVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQSWVALAVPPGE
                     EQRYTCQVEHPGLDQPLTATWEPSLSNTLVTGVISGIAVCVIIFLIGILFRILRKRQA
                     SRGAMGDYVLAECE"
BASE COUNT      279 a    363 c    397 g    280 t
ORIGIN


  Query Match             62.1%;  Score 893.6;  DB 4;  Length 1319;
  Best Local Similarity   81.7%;  Pred. No. 2.1e-252;
  Matches 1056;  Conservative    0;  Mismatches  234;  Indels    2;  Gaps    2;

Qy    137 AGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACG 196
Db      2 AGGGTGACTTCTTGGGTCACTCCTATTTCCCCGCCCCACCGAAGCAGCTGACA-CGGGCC 60

Qy    197 GGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCT 256
Db     61 AGCACCCGCGGTCCGAGCCGGGGAAATGGGCCCGCGAGCCCGACCCGCGCTCTTCTTCCT 120

Qy    257 GATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTA 316
Db    121 GATCCTCCTGCGGACCGTGGCCGCGCAGGGGCGACCGCCGCGGTCACACTCTCTGCGCTA 180

Qy    317 CCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTA 376
Db    181 CCTCTTCATGGGCGCCTCAGAGCGGGACCATGGCCTGCCCCTGTTTGAAGCCTTGGGCTA 240

Qy    377 CGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGAGCCCCGAAC 436
Db    241 TGTGGATGATGAGCTTTTCGTGGCCTACAATCACGAGAGTCGCCGTGTCAGAGTCCCGTGC 300

Qy    437 TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA 496
Db    301 CCAGTGGGTCTTGGGTGAGGCCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAA 360

Qy    497 AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAG 556
Db    361 AGGGTGGGATCACATGTTCATCGTCGACTTCTGGACCATCATGGACAACCACAACCACAG 420

Qy    557 CAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTAC 616
Db    421 CAAGGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGGTGCAAGAGGACAACAGCAC 480

Qy    617 CGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACAC 676
Db    481 CAGGGGGTTCTGGATGTACGGGTACGACGGGCAGGACCATCTTGAATTCTGCCCTGAGAC 540

Qy    677 ACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCA 736
Db    541 ACTGGATTGGAGAGCAGCAGAGTCCAGGGCCCTGACCACCAAGCTGGAGTGGGAAGTGAA 600

Qy    737 CAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCA 796
Db    601 CAAGATTCGGGCCAAGCAGAACAGAGCCTATCTCGAGAGGGACTGCCCTGAGCAGCTGCA 660

Qy    797 GCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGT 856
Db    661 GTGGTTGCTGGAGCTGGGGAGAGGGGTTCTGGACCAGCAAGTGCCTCCTTTGGTGAAGGT 720

Qy    857 GACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCC 916
```

```
Db    721 GACTCATCATGTGGCCTCTGCAGTGACCACTCTACGGTGTCAGGCTCTGAACTTCTACCC 780

Qy    917 CCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGA 976
          ||||||||||||||||| |||| ||||||| | | ||||| |||||| ||||||   |||
Db    781 CCAGAACATCACCATGAGGTGGTTGAAGGACAGGAAGCCAGTGGATGTCAAGGACGCCGA 840

Qy    977 ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGT 1036
            |||| || ||  |||||| ||||||||||||||||||| ||||| |  | ||||||||
Db    841 GTCTAAGGATGTGCTGCCCAGCGGGGATGGGACCTACCAGAGCTGGGTGGCTTTGGCTGT 900

Qy   1037 ACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCC 1096
          ||||| ||||||||||||||||||||| | ||||||||||||||||||||||||||||||
Db    901 GCCCCCCGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACCCAGGCCTGGATCAGCC 960

Qy   1097 CCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGG 1156
          ||||| ||  | |||||||||||||||  |||  |||||| |||| ||||||||||||||
Db    961 CCTCACTGCCACCTGGGAGCCCTCACTTTCTAACACCCTCGTCACTGGAGTCATCAGTGG 1020

Qy   1157 AATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAG 1216
          |||||||||| |||| |||||||  |||||||||||||||||| ||| |||||||| ||
Db   1021 GATTGCTGTTTGTGTCATCATCTTCCTCATTGGAATTTTGTTCAGAATCTTAAGGAAAAG 1080

Qy   1217 GCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGC 1276
          || | ||||||||||||||||||| |||||||||||||||| |||||||| | ||||
Db   1081 ACAAGCTTCAAGAGGAGCCATGGGGGACTACGTCTTAGCTGAATGTGAGTGAAGCACAGC 1140

Qy   1277 CTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATG 1336
          ||||||||||  |||||| |||| ||| || | |||||||| |||||| |  | |    |
Db   1141 CTGCAGACTCCTTGTGGGGAGGACACATAATTGGAGACTCAGAGAGGGTGCAC-TGGTGG 1199

Qy   1337 AGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCT 1396
            | ||||| |||||||||| |||||||||||||  |  |  |         | |   ||
Db   1200 CGTTCTTCGTGTTTCAGGACAGAGTTGAACCTAACAAATGGGAACTGTCCCAGGAGGTCT 1259

Qy   1397 TGATTTTAGCCTTCTCTGTTCATTTCCTCAAA 1428
             | ||||||||| ||||||| |  |  |||
Db   1260 CTCTCTTAGCCTTCCCTGTTCACTCTCAAAAA 1291


Search completed: June 19, 2002, 08:24:38
Job time: 6364 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 07:21:40 ; Search time 2546.54 Seconds
(without alignments)
7632.166 Million cell updates/sec

Title: US-09-497-957-12
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		735.4	51.1	819	10	BG747345	BG747345 602704818
2		630.6	43.8	1723	11	AK009581	AK009581 Mus muscu
3		505.2	35.1	570	10	BE272926	BE272926 601171213
4	c	356.4	24.8	384	10	BF883952	BF883952 PM4-ET020
5		282.6	19.6	550	10	BI339179	BI339179 364041 MA
6		275	19.1	523	10	BF080089	BF080089 230846 MA
7		257.8	17.9	489	10	BE994943	BE994943 UI-M-CG0p
8		251.4	17.5	457	9	AI850020	AI850020 UI-M-BG0-
9		248.4	17.2	455	10	BE995172	BE995172 UI-M-CG0p
10	c	230	16.0	464	9	AA217236	AA217236 mu89b05.r
11		226.4	15.7	268	10	W21141	W21141 zb51g06.r1
12		214.6	14.9	546	10	BE809138	BE809138 214520 MA
13		209.6	14.6	481	9	BB851691	BB851691 BB851691
14		202.2	14.0	831	10	BI452668	BI452668 603169877
15		197.2	13.7	502	9	BB858165	BB858165 BB858165
16		196.6	13.7	392	10	BF465475	BF465475 UI-M-CG0p
17		184.6	12.8	1208	11	BC005306	BC005306 Homo sapi
18		174	12.1	394	10	BF464345	BF464345 UI-M-CG0p
19	c	168.8	11.7	444	12	AZ025590	AZ025590 RPCI-23-3
20	c	160	11.1	334	9	AW902003	AW902003 QV0-NN102
21	c	155.2	10.8	536	12	AZ074871	AZ074871 RPCI-23-4
22		154	10.7	805	10	BG402460	BG402460 602466163
23		152	10.6	1016	9	AL532717	AL532717 AL532717
24		151.8	10.5	289	10	H33644	H33644 EST109830 R
25	c	150	10.4	481	12	AZ025784	AZ025784 RPCI-23-3
26		146	10.1	531	12	AQ703007	AQ703007 HS_5443_B
27		144	10.0	1031	9	AL513728	AL513728 AL513728
28		142	9.9	1298	11	AK005051	AK005051 Mus muscu
29		141.6	9.8	1121	9	AL545738	AL545738 AL545738
30		140	9.7	972	9	AL541209	AL541209 AL541209
31		139.8	9.7	959	9	AL560013	AL560013 AL560013
32		138.2	9.6	910	9	AL558822	AL558822 AL558822
33		137.8	9.6	1020	9	AL535151	AL535151 AL535151
34		137.6	9.6	942	9	AL539924	AL539924 AL539924
35		136.4	9.5	1094	10	BM465674	BM465674 AGENCOURT
36		132.6	9.2	932	9	AL547313	AL547313 AL547313
37		131.8	9.2	679	10	BG680802	BG680802 602628602
38		131.4	9.1	538	9	BE232472	BE232472 137492 MA
39		131	9.1	936	10	BI832514	BI832514 603082113
40		131	9.1	937	9	AL540488	AL540488 AL540488
41		131	9.1	1090	10	BM477276	BM477276 AGENCOURT
42		130	9.0	604	10	BG691825	BG691825 341215 BA
43		130	9.0	930	9	AL541347	AL541347 AL541347
44		130	9.0	979	9	AL542403	AL542403 AL542403
45		130	9.0	979	9	AL557373	AL557373 AL557373

ALIGNMENTS

RESULT 1
BG747345
LOCUS BG747345 819 bp mRNA linear EST 15-MAY-2001
DEFINITION 602704818F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4857941 5', mRNA sequence.
ACCESSION BG747345
VERSION BG747345.1 GI:14057998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 819)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1711 row: d column: 06
High quality sequence stop: 792.
FEATURES Location/Qualifiers
source 1. .819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4857941"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 202 a 201 c 235 g 181 t
ORIGIN

Query Match 51.1%; Score 735.4; DB 10; Length 819;
Best Local Similarity 95.9%; Pred. No. 3.8e-194;
Matches 787; Conservative 0; Mismatches 31; Indels 3; Gaps 3;

```
Qy    384 GACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGG 443
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      1 GACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGG 60

Qy    444 GTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG 503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG 120

Qy    504 GATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAG 563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAG 180

Qy    564 TCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGC 623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGC 240

Qy    624 TACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGAT 683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGAT 300

Qy    684 TGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATT 743
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    301 TGGAGAGCAGCAGAACCCA-GGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATT 359

Qy    744 CGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTG 803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTG 419

Qy    804 CTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACAT 863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 CTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACAT 479

Qy    864 CATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAAC 923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 CATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAAC 539

Qy    924 ATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAA 983
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    540 ATCACCATGAAGTGGCTGAAGGATAAGCAG-CAATGGATGCCAAGGAGTTCGAACCTAAA 598

Qy    984 GACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCT 1043
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    599 GACGTATTGCCCAATGGGGATGGGACCTACCA-GGCTGGATAACCTTGGCTGTACCCCCT 657

Qy   1044 GGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATT 1103
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    658 GGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATT 717

Qy   1104 GTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCT 1163
          ||||||| ||   ||    |||||||||||       | | |||||||    | | |
Db    718 GTGATCTTGGGAGCCCTCACGTCTGGCACCTAGTCATTGGAAGTCATCCAGTGGAATTGC 777

Qy   1164 GTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAAT 1204
          |||||||||||||||||| ||||||||||||||||||||
Db    778 TGTTTTGTCGTCATCTTGTTCCTTGGAATTTTGTTCATAAT 818
```

RESULT 2
AK009581
LOCUS AK009581 1723 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310032M04:hemochromatosis, full insert sequence.

ACCESSION AK009581
VERSION AK009581.1 GI:12844462
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:2310032M04.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1723)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase

and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

```
FEATURES             Location/Qualifiers
     source          1. .1723
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="MGD:MGI:1905246"
                     /db_xref="taxon:10090"
                     /clone="2310032M04"
                     /sex="male"
                     /tissue_type="tongue"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     gene            99. .1178
                     /gene="Hfe"
     CDS             99. .1178
                     /gene="Hfe"
                     /note="data source:MGD, source key:MGI:109191,
                     evidence:ISS
                     hemochromatosis
                     putative"
                     /codon_start=1
                     /protein_id="BAB26373.1"
                     /db_xref="GI:12844463"
                     /translation="MSLSAGLPVRPLLLLLLLLWSVAPQALPPRSHSLRYLFMGASEP
                     DLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWLHLSQSLKGWDYMF
                     IVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWRYGYDGQDHLEF
                     CPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLELGRGVLGQQV
                     PTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVLPNGDETY
                     QGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWEPLQSQAMIIGIISGVTVCAIFLVG
                     ILFLILRKRKASGGTMGGYVLTDCE"
     polyA_signal    1695. .1700
                     /note="putative"
     polyA_site      1723
                     /note="putative"
BASE COUNT      406 a    456 c    454 g    407 t
ORIGIN
```

```
  Query Match           43.8%;  Score 630.6;  DB 11;  Length 1723;
  Best Local Similarity   70.1%;  Pred. No. 1e-164;
  Matches  954;  Conservative    0;  Mismatches  359;  Indels   48;  Gaps    6;

Qy    105 CTGGGCATCTCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTT 164
          | ||  |  ||  || |||  ||||| |||  |||||||||||| |   | ||| |||||
Db      1 CCGGAGAGGTCTGGAACCTCAGCAATGGCTACAGGGTGACTTCTTGGATCCTCCACGTTT 60

Qy    165 CCCCGCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATG 224
          ||    ||     ||||   |  ||   | |       || || ||           |
Db     61 CCAGATCCTAGTGAAGACCGGTGGACCCAGC-------TGAGGACATGAGCCTATCAGCT 113

Qy    225 GGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAG 284
          || |  |  |   ||||| ||| || || ||| | ||  ||  | ||| || || ||||
Db    114 GGGCTCCCTGTGCGGCCGCTGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAG 173

Qy    285 GGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGAC 344
          |  |     | |||||||||| |||||    ||||||||||||||||||||||||  |||
Db    174 GCACTGCCACCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGAC 233

Qy    345 CTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTAT 404
          || || ||  | |||||||| |||  |||||| ||||||||||||||| || |||| |||
Db    234 CTCGGGCTGCCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTAC 293

Qy    405 GATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGC 464
          || ||||||||||||||  |||||| |  | || ||| | |       ||  ||||||
Db    294 AATCATGAGAGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGC 353

Qy    465 CAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGAC 524
```

```
          ||| |||||||||| |||||||||||  ||||||||||||||  |||||||||  || |||
Db    354 CAGCTGTGGCTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGAC 413

Qy    525 TTCTGGACTATTATGGAAAATCACAACCACAGCAAG---------------------------- 560
          |||||||| || |||  ||  | ||||||| |||
Db    414 TTCTGGACCATCATGGGCAACTATAACCACAGTAAGGTCACGAAGTTGGGAGTGGTGTCC 473

Qy    561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
          |||||||||| |||||||||  |||| ||||||||  |||| ||||||||||||||||
Db    474 GAGTCCCACATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGC 533

Qy    621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
          |||| |||||  || || ||||| ||||| || ||||| ||||||||||  | |||||
Db    534 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTA 593

Qy    681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
          | ||||| ||||| || ||  |||||||| |||||||| |||| |||||   |||||||
Db    594 AACTGGAGCGCAGCCGAGCCAGGGGCCTGGGCCACCAAGGTGGAATGGGACGAGCACAAG 653

Qy    741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
          || || ||||  |||||||||| |||||||||| ||||||||| |  |||||| | | |
Db    654 ATCCGTGCCAAACAGAACAGGGACTACCTGGAGAAGGACTGCCCCGAGCAGCTGAAACGG 713

Qy    801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACA 860
          | ||||||||||||||| ||| |||  || |||||||| |||||||||| |||||
Db    714 CTCCTGGAGCTGGGGAGAGGCGTTCTGGGACAGCAAGTGCCTACTTTGGTGAAAGTGACT 773

Qy    861 CATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG 920
          |  ||   | ||||| | | |||| ||||| |||||| |||  || ||| || |||||||
Db    774 CGCCACTGGGCCTCTACGGGGACCTCTCTAAGGTGTCAGGCTCTGGACTTCTTCCCCCAG 833

Qy    921 AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCT 980
          |||||||| |||| |||| |||||||  || || ||| |||||||||  ||  || | ||
Db    834 AACATCACTATGAGGTGGTTGAAGGACAACCAACCACTGGATGCCAAAGATGTCAACCCC 893

Qy    981 AAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCC 1040
          |  | ||  | || || ||||||| |||||| || |||||| | || ||||| ||  ||
Db    894 GAGAAGGTGCTACCTAACGGGGATGAGACCTATCAAGGCTGGCTGACATTGGCCGTGGCC 953

Qy   1041 CCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTC 1100
          |||||||| |||   || |  || |  || ||||||||||||||||||||| ||||| |||
Db    954 CCTGGGGACGAGACAAGGTTCACCTGTCAAGTGGAGCACCCAGGCCTGGACCAGCCTCTC 1013

Qy   1101 ATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATT 1160
          | ||     |||||||||||  |  |||    || |  | || ||| ||||||||||| |
Db   1014 ACTGCCTCTTGGGAGCCCTTGCAATCTCAGGCCATGATTATCGGAATCATCAGTGGAGTC 1073

Qy   1161 GCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAG 1220
          | || | || | || |||||    |||||||| ||||| |||| |||||||| ||| ||
Db   1074 ACCGTCTGTGCCATCTTCTTGG---TTGGAATTCTGTTCCTAATCTTAAGGAAAAGGAAG 1130

Qy   1221 GGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGC 1280
          | |||| ||||| |||||||   ||| |||||| | ||  ||||||||   |||||||||
Db   1131 GCTTCAGGAGGAACCATGGGTGGCTATGTCTTAACAGACTGTGAGTGATCTGCAGCCTGC 1190

Qy   1281 AGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCT 1340
          ||  ||| | |  || | ||  | | | |||||   || |||    | |      |
Db   1191 TGAACCACGGAAGAGAGAAAACTCAGCCAAAGACTTGGAGGGGGCACACTTGCTCCA--- 1247

Qy   1341 CTTCATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCCTTGA 1399
                |  |||| | ||||| |||||| ||| ||||| |||||||  |   |  ||
Db   1248 -------CTGTAGGACACAGTTGGACCTAACACACAGAAACTGCCTGAGAA---CTGTGC 1297

Qy   1400 TTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          | ||||| |||||||||| || || ||   | |||| |||
Db   1298 TCTTAGCTTTCTCTGTTCACTTTCTTAAGGTGTTTTCTCCA 1338
```

```
RESULT   3
BE272926
LOCUS       BE272926                 570 bp    mRNA    linear   EST 13-JUL-2000
DEFINITION  601171213F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:3544803 5',
            mRNA sequence.
```

```
ACCESSION   BE272926
VERSION     BE272926.1  GI:9147279
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM240  row: j  column: 04
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3544803"
                     /clone_lib="NIH_MGC_14"
                     /tissue_type="renal cell adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      140 a    148 c    175 g    107 t
ORIGIN


  Query Match           35.1%;  Score 505.2;  DB 10;  Length 570;
  Best Local Similarity   99.2%;  Pred. No. 6.3e-130;
  Matches  518;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy     560 GGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGA 619
           | ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      48 GCAGTCCCACACCGTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGA 107

Qy     620 GGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACT 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108 GGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACT 167

Qy     680 GGATTGGAGAGCAGCAGAACCCAG-GGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACA 738
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     168 GGATTGGAGAGCAGCAGAACCCAGTGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACA 227

Qy     739 AGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGC 798
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 AGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGC 287

Qy     799 AGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGA 858
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 AGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGA 347

Qy     859 CACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCC 918
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     348 CACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCC 407

Qy     919 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAAC 978
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db*    408 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAAC 467

Qy     979 CTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTAC 1038
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     468 CTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTAC 527

Qy    1039 CCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
           |||||||||||||||||||||||||| ||||||||||||||
Db     528 CCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 569


RESULT   4
BF883952/c
LOCUS       BF883952                384 bp    mRNA    linear   EST 17-JAN-2001
DEFINITION  PM4-ET0209-151200-003-f07 ET0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF883952
VERSION     BF883952.1  GI:12274078
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 384)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-ET0209-
            151200-003-f07&t3=2000-12-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 384.
FEATURES             Location/Qualifiers
     source          1. .384
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ET0209"
                     /dev_stage="Adult"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       92 a    112 c     87 g     93 t
ORIGIN


  Query Match           24.8%;  Score 356.4;  DB 10;  Length 384;
  Best Local Similarity   99.5%;  Pred. No. 1.9e-88;
  Matches  368;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     325 TGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATG 384
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384 TGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATG 325

Qy     385 ACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGG 444
           |||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
```

```
Db     324 ACCAGCTG-TCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGG 266

Qy     445 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGG 504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGG 206

Qy     505 ATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGT 564
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 ATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGT 146

Qy     565 CCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCT 624
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 CCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCT 86

Qy     625 ACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATT 684
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 ACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATT 26

Qy     685 GGAGAGCAGC 694
           ||||||||||
Db      25 GGAGAGCAGC 16
```

```
RESULT   5
BI339179
LOCUS        BI339179                 550 bp    mRNA    linear   EST 30-JUL-2001
DEFINITION   364041 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION    BI339179
VERSION      BI339179.1  GI:15032462
KEYWORDS     EST.
SOURCE       pig.
  ORGANISM   Sus scrofa
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE    1  (bases 1 to 550)
  AUTHORS    Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
             Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
             and Keele,J.W.
  TITLE      Design and use of two pooled tissue normalized cDNA libraries for
             EST discovery in swine
  JOURNAL    Unpublished (2000)
COMMENT      Contact: Smith TPL
             USDA, ARS, US Meat Animal Research Center
             PO Box 166, Clay Center, NE 68933-0166, USA
             Tel: 402 762 4366
             Fax: 402 762 4390
             Email: smith@email.marc.usda.gov
             Single pass sequencing. Bases called and alt_trimmed with phred
             v0.980904.e. Vector identified by cross_match with the -minscore 18
             and -minmatch 12 options.
             PCR PRimers
             FORWARD: AGGAAACAGCTATGACCAT
             BACKWARD: GTTTTCCCAGTCACGACG
             Plate: 100  row: C  column: 24
             Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .550
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 2PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT      108 a    180 c    164 g     98 t
ORIGIN
```

```
  Query Match           19.6%;  Score 282.6;  DB 10;  Length 550;
  Best Local Similarity  74.0%;  Pred. No. 9.3e-68;
  Matches  412;  Conservative    0;  Mismatches  114;  Indels   31;  Gaps    3;
```

```
Qy     140 GTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACGGGG 199
           |||||||| |||||   |   || |||  ||| ||||  |  |   ||      |  ||
Db       1 GTGACTTCTTGAGCCCCTCGGTTTCCCCGCCCCACCAACGGCCGAAAAG-----CCCTGG 55

Qy     200 ACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGAT 259
           | | |||| | |||| |||  || |||||| | |||| ||||||||||||| ||||||||
Db      56 AAGCGCGGTCCGAGCCGGG--AAAGGGCCCACAAGCCCGGCCGGCGCTTCTGCTCCTGAT 113

Qy     260 GCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCT 319
           ||  ||| ||||| || |  |||| |||   || |||  ||||||| |||| || |||
Db     114 CCTCCTGCGGACCGTGGCCACGCAGGAGCGAGCGCCGCGGCCACACTCCCTGCTCTTCCT 173

Qy     320 CTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGT 379
           |||||||| ||||| |||  || || || ||  || ||||||| || ||||||||||
Db     174 CTTCATGGGCGCCTCGGAGCCAGATCTCGGGCTGCCCCTGTTTGAGGCCTTGGGCTACGT 233

Qy     380 GGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCC 439
           ||| |||||||||| |||| |||  || | ||||||||||||  ||||| ||  | ||
Db     234 GGACGACCAGCTGTTTGTGTCCTACAATCACGAGAGTCGCCGTGCAGAGCCTCGCGCCCC 293

Qy     440 ATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGG 499
           ||| | ||  |||     || | |||  |||||||||||| || ||||| ||||||||
Db     294 CTGGCTCTCGGGTAAGGCCTCCAACCAACTGTGGCTGCAGCTAAGCCAGAGCCTGAAAGG 353

Qy     500 GTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAA 559
           |||||||||||||||||| || |||||||||| || ||||| || |||||| ||||||||
Db     354 GTGGGATCACATGTTCACCGTGGACTTCTGGACCATCATGGACAACCACAACTACAGCAA 413

Qy     560 G-----------------------GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGA 595
           |                       ||||||||||||||||||| ||||||||||||||
Db     414 GATAACGAAGCTGGGAGTGTTGCCAGAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGA 473

Qy     596 AATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCA 655
           | ||||||  |||||||| |||   || | |||||||| ||||| ||||||||||||||
Db     474 AGTGCAAGCGGACAACAGCACCAGAGGGTTCTGGAAGTATGGGTACGATGGGCAGGACCA 533

Qy     656 CCTTGAATTCTGCCCTG 672
           ||| || |||  |||||
Db     534 CCTGGAGTTCCACCCTG 550
```

```
RESULT   6
BF080089
LOCUS        BF080089                 523 bp    mRNA    linear   EST 18-OCT-2000
DEFINITION   230846 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION    BF080089
VERSION      BF080089.1  GI:10873919
KEYWORDS     EST.
SOURCE       pig.
  ORGANISM   Sus scrofa
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE    1  (bases 1 to 523)
  AUTHORS    Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
             Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
             and Keele,J.W.
  TITLE      Design and use of two pooled tissue normalized cDNA libraries for
             EST discovery in swine
  JOURNAL    Unpublished (2000)
COMMENT      Contact: Smith TPL
             USDA, ARS, US Meat Animal Research Center
             PO Box 166, Clay Center, NE 68933-0166, USA
             Tel: 402 762 4366
             Fax: 402 762 4390
             Email: smith@email.marc.usda.gov
             Single pass sequencing. Bases called and alt_trimmed with phred
             v0.980904.e. Vector identified by cross_match with the -minscore 18
             and -minmatch 12 options.
             PCR PRimers
             FORWARD: AGGAAACAGCTATGACCAT
             BACKWARD: GTTTTCCCAGTCACGACG
             Plate: 48  row: E  column: 9
             Seq primer: ATTTAGGTGACACTATAG.
```

```
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 2PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT       103 a    175 c    152 g     93 t
ORIGIN
```

```
  Query Match             19.1%;  Score 275;  DB 10;  Length 523;
  Best Local Similarity   74.0%;  Pred. No. 1.2e-65;
  Matches  387;  Conservative    0;  Mismatches  110;  Indels   26;  Gaps    2;

Qy    119 AGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAA 178
          ||||| |||||| ||| |||||||||||| |||||| | | ||||||||||||||    |
Db      2 AGCCTTGGCAATGGCTCCAGGGTGACTTCTTGAGCCACCTCGGTTTCCCCGCCCCAGCCA 61

Qy    179 AGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAG 238
          ||  || || |   | | || |     | ||    |  |||||||||||| | |||| |
Db     62 AGCGGCCGAAAAGCACCTGGAAGCGCGGTCCGAGCC--GGGAAATGGGCCCACAAGCCCG 119

Qy    239 GCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCG 298
          |||||||||||| |||||||| ||  ||| ||||| || |  ||| |  ||   || |||
Db    120 GCCGGCGCTTCTGCTCCTGATCCTCCTGCGGACCGTGGCCACGCATGATCGAGCGCCGCG 179

Qy    299 TTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTT 358
          |||||||  |||| || |||||||||||| ||||| ||||  || || || ||  || |
Db    180 GCCACACTCCCTGCTCTTCCTCTTCATGGGCGCCTCGGAGCCAGATCTCGGGCTGCCCCT 239

Qy    359 GTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCG 418
          ||||||  || |||||||||||||| |||||||||| |||| |||  ||  | ||||||||
Db    240 GTTTGAGGCCTTGGGCTACGTGGACGACCAGCTGTTTGTGTCCTACAATCACGAGAGTCG 299

Qy    419 CCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCA 478
          |||||  ||||| ||  | || ||| | ||  |||     || | |||  ||||||||||
Db    300 CCGTGCAGAGCCTCGCGCCCCCTGGCTCTCGGGTAAGGCCTCCAACCAACTGTGGCTGCA 359

Qy    479 GCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTAT 538
          ||| || ||||| |||||||||||||||||||||||||| || ||||||||||| || ||
Db    360 GCTAAGCCAGAGCCTGAAAGGGTGGGATCACATGTTCACCGTGGACTTCTGGACCATCAT 419

Qy    539 GGAAAATCACAACCACAGCAAG-----------------------GAGTCCCACACCCT 574
          ||| || |||||| ||||||||                       ||||||||||||||
Db    420 GGACAACCACAACTACAGCAAGATAACGAAGCTGGGAGTGTGGCCAGAGTCCCACACCCT 479

Qy    575 GCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACC 617
          |||||| |||||||||||||| ||||||  |||||||| |||
Db    480 GCAGGTGATCCTGGGCTGTGAAGTGCAAGCGGACAACAGCACC 522
```

```
RESULT   7
BE994943
LOCUS       BE994943                489 bp    mRNA    linear   EST 05-OCT-2000
DEFINITION  UI-M-CG0p-bik-d-03-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
            UI-M-CG0p-bik-d-03-0-UI 3', mRNA sequence.
ACCESSION   BE994943
VERSION     BE994943.1  GI:10678689
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
```

```
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
            clones from RESEARCH GENETICS. It should be noted that Bento Soares
            is generating a small number of additional specialized
            non-redundant arrays of BMAP cDNAs whose availability will be
            considered under appropriate and limited collaborative arrangements
            The following repetitive elements were found in this cDNA sequence:
            1-31, >(CAG)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-CG0p-bik-d-03-0-UI"
                     /clone_lib="NIH_BMAP_Ret4_S2"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_Ret4_S2 library is a subtracted library,
                     ultimately derived from mouse retina tissue libraries at
                     various stages of development. For a detailed description
                     of the library from which this clone was derived,  please
                     visit our web site at brainest.eng.uiowa.edu.
                     TAG_SEQ=None found"
BASE COUNT       103 a    133 c    144 g    109 t
ORIGIN
```

```
  Query Match             17.9%;  Score 257.8;  DB 10;  Length 489;
  Best Local Similarity   73.1%;  Pred. No. 7.5e-61;
  Matches  356;  Conservative    0;  Mismatches  107;  Indels   24;  Gaps    1;

Qy    238 GGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGC 297
          |||||  ||| || || ||| | ||  ||  | ||| || || |||||  |     | ||
Db      3 GGCCGCTGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAGGCACTGCCACCGC 62

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          |||||||  |||||     |||||||||||||||||||||||||  |||||  || ||  | |
Db     63 GTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGACCTCGGGCTGCCTT 122

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
          |||||||  |||  |||||| |||||||||||||| || |||| |||  ||  |||||||||
Db    123 TGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATCATGAGAGTC 182

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||  ||||||  |  | || ||| | |        ||  ||||||||| |||||||||
Db    183 GCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGCTGTGGCTGC 242

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          | |||||||||| |||||||||||||  |||||||||  || ||||||||||| || |
Db    243 ATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTCTGGACCATCA 302

Qy    538 TGGAAAATCACAACCACAGCAAG-----------------------GAGTCCCACACCC 573
          |||  ||  | |||||||| |||                       |||||||||| ||
Db    303 TGGGCAACTATAACCACAGTAAGGTCACGAAGTTGGGAGTGGTGTCCGAGTCCCACATCC 362

Qy    574 TGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGT 633
          |||||||  |||| ||||||||  |||| ||||||||||||||   |||| |||||  |
Db    363 TGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGCGGCTTCTGGAGAT 422

Qy    634 ACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAG 693
          | || ||||| ||||| || ||||| ||||||||||  | |||||  | ||||| ||||
Db    423 ATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTAAACTGGAGCGCAG 482
```

```
Qy     694 CAGAACC 700
           | || ||
Db     483 CCGAGCC 489


RESULT   8
AI850020
LOCUS       AI850020                457 bp    mRNA    linear   EST 15-JUL-1999
DEFINITION  UI-M-BG0-aib-g-10-0-UI.s1 NIH_BMAP_MSC Mus musculus cDNA clone
            UI-M-BG0-aib-g-10-0-UI 3', mRNA sequence.
ACCESSION   AI850020
VERSION     AI850020.1  GI:5493926
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
            available by the means that is soon to be determined. When NIH
            determines the means for distribution of the BMAP cDNA clones, this
            record will be updated accordingly when that means is determined.
            The following repetitive elements were found in this cDNA sequence:
            3-30, >(CAG)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BG0-aib-g-10-0-UI"
                     /clone_lib="NIH_BMAP_MSC"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_MSC library is a non-normalized library
                     constructed from  mouse spinal cord. The tag is a string
                     of 5 nucleotides present  between the Not I site and the
                     oligo-dT track.  The library was  constructed as described
                     by Bonaldo, Lennon and Soares,  Genome Research 6: 791-806
                     , 1996. Tissue provided by Ms. Annie Novakovich,
                     Zivic-Miller Laboratories.
                     TAG_LIB=NIH_BMAP_MSC
                     TAG_TISSUE=spinal-cord
                     TAG_SEQ=TCGAA"
BASE COUNT      100 a    129 c    144 g     84 t
ORIGIN


  Query Match            17.5%;  Score 251.4;  DB 9;  Length 457;
  Best Local Similarity  79.6%;  Pred. No. 4.4e-59;
  Matches  297;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy     561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
           |||||||||| |||||||||  |||| ||||||||  |||| ||||||||||||||||
Db      85 GAGTCCCACATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGC 144

Qy     621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
           |||| |||||  || || ||||| ||||| || ||||| ||||||||||  | |||||
Db     145 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTA 204

Qy     681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
           | ||||| ||||| || ||  |||||||| |||||||| |||| |||||   |||||||
Db     205 AACTGGAGCGCAGCCGAGCCAGGGGCCTGGGCCACCAAGGTGGAATGGGACGAGCACAAG 264

Qy     741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
           || || |||| |||||||||| |||||||||| ||||||||| |  |||||| | | |
Db     265 ATCCGTGCCAAACAGAACAGGGACTACCTGGAGAAGGACTGCCCCGAGCAGCTGAAACGG 324

Qy     801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACA 860
           | |||||||||||||||| ||| |||  || |||||||| |||||||||| |||||
Db     325 CTCCTGGAGCTGGGGAGAGGCGTTCTGGGACAGCAAGTGCCTACTTTGGTGAAAGTGACT 384

Qy     861 CATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG 920
           |  ||   | ||||| | | |||| ||||| |||||| |||  || ||| || |||||||
Db     385 CGCCACTGGGCCTCTACGGGGACCTCTCTAAGGTGTCAGGCTCTGGACTTCTTCCCCCAG 444

Qy     921 AACATCACCATGA 933
           |||||||| ||||
Db     445 AACATCACTATGA 457


RESULT   9
BE995172
LOCUS       BE995172                455 bp    mRNA    linear   EST 05-OCT-2000
DEFINITION  UI-M-CG0p-bil-h-10-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
            UI-M-CG0p-bil-h-10-0-UI 3', mRNA sequence.
ACCESSION   BE995172
VERSION     BE995172.1  GI:10679153
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
            clones from RESEARCH GENETICS. It should be noted that Bento Soares
            is generating a small number of additional specialized
            non-redundant arrays of BMAP cDNAs whose availability will be
            considered under appropriate and limited collaborative arrangements
            The following repetitive elements were found in this cDNA sequence:
            1-31, >(CAG)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-CG0p-bil-h-10-0-UI"
                     /clone_lib="NIH_BMAP_Ret4_S2"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_Ret4_S2 library is a subtracted library,
```

ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu.
TAG_SEQ=None found"
BASE COUNT 99 a 129 c 144 g 83 t
ORIGIN

Query Match 17.2%; Score 248.4; DB 10; Length 455;
Best Local Similarity 79.5%; Pred. No. 3e-58;
Matches 294; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```
Qy    561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
          |||||||||| |||||||||  |||| ||||||||  |||| ||||||||||||||||
Db     86 GAGTCCCACATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGC 145

Qy    621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
          |||| |||||  || || ||||| ||||| || ||||| ||||||||||  | |||||
Db    146 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTA 205

Qy    681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
          | ||||| ||||| || ||  |||||||| |||||||| |||| |||||   |||||||
Db    206 AACTGGAGCGCAGCCGAGCCAGGGGCCTGGGCCACCAAGGTGGAATGGGACGAGCACAAG 265

Qy    741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
          || || ||||  |||||||||| |||||||||| |||||||| |  |||||| | | |
Db    266 ATCCGTGCCAAACAGAACAGGGACTACCTGGAGAAGGACTGCCCCGAGCAGCTGAAACGG 325

Qy    801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACA 860
          | |||||||||||||||| ||| |||  || |||||||| |||||||||| |||||
Db    326 CTCCTGGAGCTGGGGAGAGGCGTTCTGGGACAGCAAGTGCCTACTTTGGTGAAAGTGACT 385

Qy    861 CATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG 920
          |  ||   | ||||| | | |||| ||||| |||||| |||  || ||| || |||||||
Db    386 CGCCACTGGGCCTCTACGGGGACCTCTCTAAGGTGTCAGGCTCTGGACTTCTTCCCCCAG 445

Qy    921 AACATCACCA 930
          |||||||| |
Db    446 AACATCACTA 455
```

RESULT 10
AA217236/c
LOCUS AA217236 464 bp mRNA linear EST 06-FEB-1997
DEFINITION mu89b05.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone IMAGE:652689 5' similar to TR:G940354 G940354 CLASS I HISTOCOMPATIBILITY ANTIGEN-LIKE PROTEIN. ;, mRNA sequence.
ACCESSION AA217236
VERSION AA217236.1 GI:1826237
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 464)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:398537
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .464
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:652689"
/clone_lib="Soares mouse lymph node NbMLN"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 101 a 136 c 119 g 108 t
ORIGIN

Query Match 16.0%; Score 230; DB 9; Length 464;
Best Local Similarity 74.1%; Pred. No. 4.2e-53;
Matches 329; Conservative 0; Mismatches 90; Indels 25; Gaps 2;

```
Qy    294 CTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTT 353
          | ||||||||| |||||    |||||||||||||||||||||||||  ||||| || ||
Db    453 CCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGACCTCGGGCTG 394

Qy    354 TCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAG 413
          | |||||||| |||  |||||| |||||||||||||| || |||| |||  || |||||
Db    393 CCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATCATGAG 334

Qy    414 AGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGG 473
          ||||||||||  |||||| |  | || ||| | |        ||  ||||||||| |||||
Db    333 AGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGCTGTGG 274

Qy    474 CTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACT 533
          ||||| |||||||||||  |||||||||||||  |||||||||  || ||||||||||
Db    273 CTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTCTGGACC 214

Qy    534 ATTATGGAAAATCACAACCACAGCAAG----------------------GAGTCCCAC 569
          || ||||  ||  | |||||||| |||                      |||||||||
Db    213 ATCATGGGCAACTATAACCACAGTAAGGTCACGAAGTTGGGAGTGGTGTCCGAGTCCCAC 154

Qy    570 ACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGG 629
          | |||||||||  |||| ||||||||  |||| ||||||||||||||  |||| ||||
Db    153 ATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGCGGCTTCTGG 94

Qy    630 AAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGA 689
          |  || || ||||| ||||| || ||||| ||||||||||  | |||||  | |||||
Db     93 AGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTAAACTGGAG- 33

Qy    690 GCAGCAGAACCCAGGGCCTGGCCC 713
          ||||| || ||  || |  ||| |
Db     34 GCAGCCGAGCCAGGGCCTGGGCAC 11
```

RESULT 11
W21141
LOCUS W21141 268 bp mRNA linear EST 20-AUG-1996
DEFINITION zb51g06.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:307162 5', mRNA sequence.
ACCESSION W21141
VERSION W21141.1 GI:1298018
KEYWORDS EST.
SOURCE human.

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 268)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1642   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 173.
FEATURES             Location/Qualifiers
     source          1. .268
                     /organism="Homo sapiens"
                     /db_xref="GDB:1250574"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:307162"
                     /clone_lib="Soares_fetal_lung_NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library, Soares
                     fetal heart NbHH19W."
BASE COUNT       79 a     55 c     64 g     69 t      1 others
ORIGIN
```

```
  Query Match            15.7%;  Score 226.4;  DB 10;  Length 268;
  Best Local Similarity  99.2%;  Pred. No. 3.3e-52;
  Matches  238;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      1 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG-ACTACGTCTTAGCTGAAC 59

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 119

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 179

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    180 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAAATTTCCCCA 239
```

```
RESULT  12
BE809138
LOCUS       BE809138                546 bp    mRNA    linear   EST 25-APR-2001
```

```
DEFINITION  214520 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE809138
VERSION     BE809138.1  GI:10240250
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 71  row: A  column: 5
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT      126 a    139 c    180 g    100 t      1 others
ORIGIN
```

```
  Query Match            14.9%;  Score 214.6;  DB 10;  Length 546;
  Best Local Similarity  85.9%;  Pred. No. 9.2e-49;
  Matches  238;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy    561 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 620
          |||||||||||||||||| ||||||||||||||  ||||||| |||||||| |||
Db    139 GAGTCCCACACCCTGCAGGTGATCCTGGGCTGTGAGTTGCAAGAGGACAACAGCACCAGA 198

Qy    621 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 680
          || | ||||||||||||||| |||||||||||||| ||||||||| | ||||| ||||||
Db    199 GGGTTCTGGAAGTACGGGTACGATGGGCAGGACCATCTTGAATTCCGGCCTGAGACACTG 258

Qy    681 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
          |||||||||||||||| |||||||||     ||||||||||||||||||  | |||||
Db    259 GATTGGAGAGCAGCAGAGCCCAGGGCCCAAGTCACCAAGCTGGAGTGGGAAGTGAACAAG 318

Qy    741 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 800
          ||||||||| |||||||||||||||||||  |||| ||||| |  ||||||| |||
Db    319 ATTCGGGCCAAGCAGAACAGGGCCTACCTGGATCGGGATTGCCCCGAGCAGCTGCTGCAT 378

Qy    801 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          ||||||||||||||||||   | |||| || ||||
Db    379 TTGCTGGAGCTGGGGAGAGGGCCTCTGGAGCAGCAAG 415
```

```
RESULT 13
BB851691
LOCUS       BB851691                481 bp    mRNA    linear   EST 26-NOV-2001
DEFINITION  BB851691 RIKEN full-length enriched, B16 F10Y cells Mus musculus
            cDNA clone G370002P09 5', mRNA sequence.
ACCESSION   BB851691
VERSION     BB851691.1  GI:17093145
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .481
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G370002P09"
                     /clone_lib="RIKEN full-length enriched, B16 F10Y cells"
                     /cell_type="B16 F10Y cells"
BASE COUNT       95 a    134 c    140 g    112 t
ORIGIN
```

Query Match 14.6%; Score 209.6; DB 9; Length 481;
Best Local Similarity 68.5%; Pred. No. 2.1e-47;
Matches 307; Conservative 0; Mismatches 134; Indels 7; Gaps 1;

```
Qy    114 TCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCC 173
          ||  || |||  ||||| |||  |||||||||||| |   | ||| |||||||    ||
Db      9 TCTGGAACCTCAGCAATGGCTACAGGGTGACTTCTTGGATCCTCCACGTTTCCAGATCCT 68

Qy    174 CCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGA 233
          ||||   |  ||   | |       || || ||          |   || | |
Db     69 AGTGAAGACCGGTGGACCCAGC-------TGAGGACATGAGCCTATCAGCTGGGCTCCCT 121

Qy    234 GCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTG 293
          |  |||||  ||| || || ||| | ||  ||  | ||| || || |||||  |
Db    122 GTGCGGCCGCTGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAGGCACTGCCA 181

Qy    294 CTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTT 353
          | |||||||| |||||    |||||||||||||||||||||||||  |||||  || ||
Db    182 CCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGACCTCGGGCTG 241

Qy    354 TCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAG 413
          | |||||||| |||  |||||| |||||||||||||| || |||| |||  || |||||
Db    242 CCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATCATGAG 301

Qy    414 AGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGG 473
          ||||||||||  |||||| |  | || ||| | |       ||  |||||||| |||||
Db    302 AGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGCTGTGG 361

Qy    474 CTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACT 533
          ||||| |||||||||| ||||||||||||||  |||||||||  || ||||||||||||
Db    362 CTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTCTGGACC 421

Qy    534 ATTATGGAAAATCACAACCACAGCAAGG 561
          || ||||  ||  | |||||||| ||||
Db    422 ATCATGGGCAACTATAACCACAGTAAGG 449
```

```
RESULT 14
BI452668
LOCUS       BI452668                831 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603169877F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5249395 5',
            mRNA sequence.
ACCESSION   BI452668
VERSION     BI452668.1  GI:15243324
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11629  row: j  column: 20
            High quality sequence stop: 818.
FEATURES             Location/Qualifiers
     source          1. .831
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5249395"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT       207 a    220 c    189 g    215 t
ORIGIN
```

```
  Query Match           14.0%;  Score 202.2;  DB 10;  Length 831;
  Best Local Similarity  68.3%;  Pred. No. 3.3e-45;
  Matches  345;  Conservative    0;  Mismatches  143;  Indels   17;  Gaps    4;

Qy     935 GTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCC 994
           |||| ||||||| || || ||| |||||||||| ||  || | ||  |  | ||  | ||
Db       2 GTGGTTGAAGGACAACCAACCACTGGATGCCAAAGATGTCAACCCCGAGAAGGTGCTACC 61

Qy     995 CAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCA 1054
           || ||||||| |||||| || |||||| | || ||||| ||  |||||||||| |||
Db      62 TAACGGGGATGAGACCTATCAAGGCTGGCTGACATTGGCCGTGGCCCCTGGGGACGAGAC 121

Qy    1055 GAGATATACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGA 1114
           || |  || |  || |||||||||||||||||| ||||| |||| ||     |||||
Db     122 AAGGTTCACCTGTCAAGTGGAGCACCCAGGCCTGGACCAGCCTCTCACTGCCTCTTGGGA 181

Qy    1115 GCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGT 1174
           |||||  |  |||    || |  | |||||| ||||||||||  |  |  | | |  |
Db     182 GCCCTTGCAATCTCAGGCCATGATTATTGGAATCATCAGTGGGGTCACCATCTGT---GC 238

Qy    1175 CATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGC 1234
           |||||| ||  |||||||| ||||| |||| |||||||| ||| ||| |||| ||||| |
Db     239 CATCTTCTTAGTTGGAATTCTGTTCCTAATCTTAAGGAAAAGGAAGGCTTCAGGAGGAAC 298

Qy    1235 CATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGG 1294
           ||||||   ||| |||||| | ||  ||||||||   |||||||| ||  ||| |  |
Db     299 CATGGGTGGCTATGTCTTAACAGACTGTGAGTGATCTGCAGCCTGCTGAACCACGGAAGA 358

Qy    1295 AAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGG 1354
           || | ||  | | | |||||   || |||    | |     |          |  |||
Db     359 GAGAAAACTCAGCCAAAGACTTGGAGGGGGCACACTTGCTCCA----------CTGTAGG 408

Qy    1355 AGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCT 1413
           | | ||||| |||||| ||| ||||| |||||||  |   |  || | ||||||||||||
Db     409 ACACAGTTGGACCTAACACACAGAAACTGCCTGAGAA---CTGTGCTCTTAGCCTTCTCT 465

Qy    1414 GTTCATTTCCTCAAAAAGATTTCCC 1438
           |||||  ||  ||  ||   ||||  ||
Db     466 GTTCACTTTCTTAAGGTGATTCTCC 490


RESULT  15
BB858165
LOCUS       BB858165                 502 bp    mRNA    linear   EST 26-NOV-2001
DEFINITION  BB858165 RIKEN full-length enriched, B16 F10Y cells Mus musculus
            cDNA clone G370049A15 5', mRNA sequence.
ACCESSION   BB858165
VERSION     BB858165.1  GI:17099619
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 502)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .502
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G370049A15"
                     /clone_lib="RIKEN full-length enriched, B16 F10Y cells"
                     /cell_type="B16 F10Y cells"
BASE COUNT       98 a    139 c    147 g    118 t
ORIGIN


  Query Match           13.7%;  Score 197.2;  DB 9;  Length 502;
  Best Local Similarity  67.8%;  Pred. No. 6.4e-44;
  Matches  308;  Conservative    0;  Mismatches  138;  Indels    8;  Gaps    2;

Qy     109 GCATCTCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCC 168
           | |  ||  || |||  ||||| |||  |||||||||||| |   | ||| |||||||
Db       3 GGAGGTCTGGAACCTCAGCAATGGCTACAGGGTGACTTCTTGGATCCTCCACGTTTCCAG 62

Qy     169 GCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCC 228
             ||    ||||   |  ||   | |       || || ||         |   || |
Db      63 ATCCTAGTGAAGACCGGTGGACCCAGC-------TGAGGACATGAGCCTATCAGCTGGGC 115

Qy     229 CGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGC 288
           |  |  ||||| ||| || || ||| | ||  ||  | ||| || || |||||  |
Db     116 TCCCTGTGCGGCCGCTGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAGGCAC 175

Qy     289 GCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTG 348
             | ||||||||| |||||    ||||||||||||||||||||| |||  ||||| |
Db     176 TGCCACCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCACAGCCAGACCTCG 235

Qy     349 GTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATG 408
           | ||  | |||||||| |||  |||||| |||||||||||||| || |||| |||  ||
Db     236 GGCTGCCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTACAATC 295

Qy     409 ATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGA 468
           ||||||||||||||  ||||||  |  | || ||| | |        ||  |||||||||
Db     296 ATGAGAGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGCCAGC 355

Qy     469 TGTGG-CTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTC 527
           |||||  ||||| |||||||||||  ||||||||||||||  |||||||||  || ||||||
Db     356 TGTGGCCTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGACTTC 415

Qy     528 TGGACTATTATGGAAAATCACAACCACAGCAAGG 561
           |||||  || ||||  ||  | |||||||| ||||
Db     416 TGGACCATCATGGGCAACTATAACCACAGTAAGG 449
```

Search completed: June 19, 2002, 07:21:51
Job time: 2597 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:49:52 ; Search time 352.95 Seconds
(without alignments)
7004.836 Million cell updates/sec

Title: US-09-497-957-12
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1440	100.0	1440	22	AAC68432	Human hereditary h
2	1438.4	99.9	1440	22	AAC68430	Human hereditary h
3	1438.4	99.9	1440	22	AAC68431	Human hereditary h
4	1436.8	99.8	1440	18	AAT96691	Hereditary haemoch
5	1436.8	99.8	1440	22	AAC68429	Human hereditary h
6	1436.8	99.8	2727	19	AAV23525	Haemochromatosis g
7	1215.8	84.4	2506	21	AAA96769	cDNA sequence enco
8	557.6	38.7	596	22	AAI63897	Human polynucleoti
9	321	22.3	10825	22	AAC68427	Human hereditary h
10	321	22.3	10825	22	AAC68428	Human hereditary h
11	319.4	22.2	10825	18	AAT96690	Hereditary haemoch
12	319.4	22.2	10825	22	AAC68425	Human hereditary h
13	319.4	22.2	10825	22	AAC68426	Human hereditary h
14	319.4	22.2	12146	21	AAA96794	Genomic DNA of a h
c 15	319.4	22.2	237326	19	AAV57903	Hereditary haemoch
16	280	19.4	517	22	AAC68441	Human hereditary h
17	278.4	19.3	359	20	AAX16055	Hereditary hemochr
18	278.4	19.3	517	22	AAC68440	Human hereditary h
19	278.4	19.3	5749	22	AAL36747	Human musculoskele
20	191.8	13.3	8622	24	ABL34142	Human immune syste
c 21	178.2	12.4	8622	24	ABL34143	Human immune syste
22	173.6	12.1	1112	21	AAA48668	cDNA encoding chic
23	172	11.9	1230	21	AAA48673	cDNA encoding chic
24	168.8	11.7	1195	21	AAA48671	cDNA encoding chic
25	166.8	11.6	1197	21	AAA48672	cDNA encoding chic
26	166.8	11.6	1262	21	AAA48691	Consensus DNA of c
27	164	11.4	1173	21	AAC78071	Human cancer assoc
28	164	11.4	1173	22	AAF72767	Human prostate can
29	162	11.2	1145	21	AAA48667	cDNA encoding chic
30	160.4	11.1	1230	21	AAA48669	cDNA encoding chic
31	160.4	11.1	1284	9	AAN80603	Probe F10 of Major
32	158.8	11.0	1230	21	AAA48665	cDNA encoding chic
33	158.8	11.0	1230	21	AAA48670	cDNA encoding chic
34	157.2	10.9	1230	21	AAA48666	cDNA encoding chic
35	155.6	10.8	1230	21	AAA48664	cDNA encoding chic
36	143.6	10.0	1554	22	AAI93004	Human polynucleoti
37	143.2	9.9	1101	12	AAQ12116	HLA-C exon Cb-1.
38	142	9.9	1567	22	AAH98676	Human EST-derived
39	141.8	9.8	1073	22	AAH42223	Nucleotide sequenc
40	141.2	9.8	1098	22	AAS07697	Human cDNA encodin
41	140	9.7	1101	12	AAQ12117	HLA-C exon Cb-2.
42	139.4	9.7	2034	23	AAS90913	DNA encoding novel
43	139.4	9.7	2037	23	AAS90740	DNA encoding novel
44	138.4	9.6	4965	16	AAQ75973	pHLA-B7/beta-2 mic
45	137.8	9.6	1089	22	AAH45555	Human cancer cell

ALIGNMENTS

RESULT 1
AAC68432
ID AAC68432 standard; DNA; 1440 BP.
XX
AC AAC68432;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1/2 mutation cDNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX

PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 348 A; 354 C; 407 G; 331 T; 0 other;

Query Match 100.0%; Score 1440; DB 22; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtcgcc 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ccttgaactactaccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtaccaggtggagcacc 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

RESULT 2
AAC68430
ID AAC68430 standard; DNA; 1440 BP.
XX
AC AAC68430;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1 mutation cDNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;

```
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 1440 BP; 348 A; 355 C; 406 G; 331 T; 0 other;


  Query Match             99.9%;  Score 1438.4;  DB 22;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccccaaaag 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtaccaggtggagcacc 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 caggcctggatcagccccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440


RESULT   3
AAC68431
ID   AAC68431 standard; DNA; 1440 BP.
XX
AC   AAC68431;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation cDNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ss.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
```

```
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 1440 BP; 347 A; 354 C; 408 G; 331 T; 0 other;
```

```
  Query Match             99.9%;  Score 1438.4;  DB 22;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtcgcc 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    1021 ggataaccttggctgtaccccctggggaagagcagagatatacgtgccaggtggagcacc 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

```
RESULT   4
AAT96691
ID   AAT96691 standard; cDNA; 1440 BP.
XX
AC   AAT96691;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene cDNA clone.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             222.1268
FT                   /*tag=  a
FT   mutation        408
FT                   /*tag=  g
FT                   /note= "C to G substitution (24d2 mutation)
FT                          results in His to Asp substitution"
FT   variation       414
```

```
FT                   /*tag=  h
FT                   /note= "A to T substitution (24d7 variant)
FT                          results in Ser to Cys substitution"
FT   mutation        1066
FT                   /*tag=  i
FT                   /note= "G to A substitution (24d1 mutation
FT                          associated with HH), results in Cys to
FT                          Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   P-PSDB; AAW36499.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 4; 115pp; English.
XX
CC   This cDNA clone, designated cDNA24, is derived from human gene
CC   whose mutated form is associated with hereditary haemochromatosis
CC   (HH).  It was obtained from a directionally cloned plasmid-based
CC   cDNA library following identification of the HH locus in the HLA
CC   region of chromosome 6.  A single mutation (24d1) in the HH gene
CC   appears responsible for the majority of HH disease.  This comprises
CC   a G to A substitution that is present in 86% of affected
CC   chromosomes and in 4% of unaffected chromosomes.  It results in a
CC   Cys to Tyr substitution in the encoded protein (see AAW36499) at a
CC   critical disulphide bridge important for secondary structure.  The
CC   following are claimed: a 10825 bp genomic DNA sequence (I) (see
CC   AAT96690), the 1437 bp cDNA sequence (Ia) and their 24d1, 24d2 and
CC   24d7 variants; a cloning or expression vector; host cells; a
CC   peptide product chosen from the HH gene product, its variants
CC   (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC   residues of these; an antibody produced using the peptide; a method
CC   to determine the presence or absence of the common HH gene
CC   mutation; an animal model for the HH disease; metal chelation
CC   agents, T-cell differentiation factors and therapeutic agents for
CC   the mitigation of injury due to oxidative process in vivo or
CC   mitigation of iron overload; a method for screening potential
CC   therapeutic agents for activity in connection with HH disease; an
CC   antisense oligonucleotide directed against a transcriptional
CC   product of a nucleic acid sequence as above; and oligonucleotides
CC   or pairs of oligonucleotides covering a range of nucleotides from
CC   (I), (Ia) or their variants, useful for detecting a polymorphism in
CC   the HH gene.  The invention also relates to methods for screening
CC   for HH homozygotes, to HH diagnosis, prenatal screening and
CC   diagnosis, and therapies of HH disease, including gene therapy,
CC   protein- and antibody-based therapeutics, and small molecule
CC   therapeutics.
XX
SQ   Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;
```

Query Match 99.8%; Score 1436.8; DB 18; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db       1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
Db      61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
Db     121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgccccccaaaag 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
Db     181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
Db     241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db     301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
Db     361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db     421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db     481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db     541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
Db     601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
Db     661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db     721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db     781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
Db     841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
Db     901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
Db     961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
Db    1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtgccaggtggagcacc 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db    1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140
```

```
Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

RESULT 5
AAC68429
ID AAC68429 standard; DNA; 1440 BP.
XX
AC AAC68429;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis cDNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;

Query Match 99.8%; Score 1436.8; DB 22; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ccttgaactactaccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtgccaggtggagcacc 1080
```

```
Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 caggcctggatcagccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

RESULT 6
AAV23525
ID AAV23525 standard; mRNA; 2727 BP.
XX
AC AAV23525;
XX
DT 10-JUL-1998 (first entry)
XX
DE Haemochromatosis gene.
XX
KW Hereditary haemochromatosis; HC gene; HH identification; diagnosis;
KW autosomal recessive disorder; ss.
XX
OS Homo sapiens.
XX
PN WO9807884-A1.
XX
PD 26-FEB-1998.
XX
PF 22-AUG-1997; 97WO-AU00539.
XX
PR 03-SEP-1996; 96AU-0002083.
PR 23-AUG-1996; 96AU-0001849.
XX
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
PI Busfield F, Cullen LM, Jazwinska EC, Powell LW;
XX
DR WPI; 1998-179064/16.
XX
PT Detection of autosomal recessive disorder - particularly hereditary
PT haemochromatosis, by detecting a mutation in the HC gene
XX
PS Disclosure; Page -; 32pp; English.
XX
CC This sequence represents the haemochromatosis (HC) gene. Mutations in
CC this sequence are detected using the method of the invention. The method
CC is for identifying an individual with hereditary haemochromatosis (HH) or
CC a predisposition to develop HH or to genetically pass on HH to an
CC offspring, comprising isolating a biological sample and amplifying a
CC region of genomic DNA in the biological sample encompassing all or part
CC of the DNA between markers D6S265 and D6S276, and detecting at least one
CC homozygous or heterozygous mutation in a nucleotide within the region.
CC The method can also be used for identifying an individual with an
CC autosomal recessive disorder (ARD) or predisposition to develop and/or
CC genetically pass on an ARD to an offspring, comprises isolating a
CC biological sample from the individual and screening genomic DNA in the
CC sample for the presence of a homozygous or heterozygous mutation in a
CC gene, the normal function of which, is required to prevent progression of
CC the disorder. The method(s) can be used to identify individuals that are
CC homozygous or heterozygous (carriers) for the mutation causing the ARD.
CC Especially the method is used to diagnose HH or predisposition to HH by
CC detecting a Cys282Tyr substitution. Individuals homozygous for this
CC mutation have HH and heterozygotes are potential carriers of the
CC disease.
XX
SQ Sequence 2727 BP; 702 A; 606 C; 660 G; 759 T; 0 other;

Query Match 99.8%; Score 1436.8; DB 19; Length 2727;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggggacactggatcacctagtgtttcacaagcaggtaccttctgctgtaggagagagaga 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 actaaagttctgaaagacctgttgcttttcaccaggaagttttactgggcatctcctgag 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cctaggcaatagctgtagggtgacttctggagccatccccgtttccccgcccccaaaag 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aagcggagatttaacggggacgtgcggccagagctggggaaatgggcccgcgagccaggc 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cggcgcttctcctcctgatgcttttgcagaccgcggtcctgcaggggcgcttgctgcgtt 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttccttgt 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 ttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtcgcc 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 tgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactattatgg 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aaaatcacaaccacagcaaggagtcccacaccctgcaggtcatcctgggctgtgaaatgc 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aagaagacaacagtaccgagggctactggaagtacgggtatgatgggcaggaccaccttg 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 aattctgccctgacacactggattggagagcagcagaacccagggcctggcccaccaagc 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 tggagtgggaaaggcacaagattcgggccaggcagaacagggcctacctggagagggact 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gccctgcacagctgcagcagttgctggagctggggagaggtgttttggaccaacaagtgc 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    841 ctcctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ccttgaactactacccccagaacatcaccatgaagtggctgaaggataagcagccaatgg 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 atgccaaggagttcgaacctaaagacgtattgcccaatggggatgggacctaccagggct 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1021 ggataaccttggctgtacccccctggggaagagcagagatatacgtgccaggtggagcacc 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 caggcctggatcagcccctcattgtgatctgggagccctcaccgtctggcaccctagtca 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttggagtcatcagtggaattgctgtttttgtcgtcatcttgttcattggaattttgttca 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 taatattaaggaagaggcagggttcaagaggagccatggggcactacgtcttagctgaac 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaactagagactcaaag 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 agggagtgcatttatgagctcttcatgtttcaggagagagttgaacctaaacatagaaat 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgcctgacgaactccttgattttagccttctctgttcatttcctcaaaaagatttcccca 1440
```

```
RESULT   7
AAA96769
ID   AAA96769 standard; cDNA; 2506 BP.
XX
AC   AAA96769;
XX
DT   19-FEB-2001  (first entry)
XX
DE   cDNA sequence encoding a histocompatibility iron loading (HFE) protein.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1044
FT                   /*tag=  a
FT                   /product= "histocompatibility iron loading (HFE) protein"
FT   sig_peptide     1..66
FT                   /*tag=  b
FT   mutation        187
FT                   /*tag=  c
FT                   /note= "if this base is mutated to G, then the
FT                          protein contains the mutation H63D"
FT   mutation        193
FT                   /*tag=  d
FT                   /note= "if this base is mutated to T, then the
FT                          protein contains the mutation S65C"
FT   mutation        277
FT                   /*tag=  e
FT                   /note= "if this base is mutated to C, then the
FT                          protein contains the mutation G93R"
FT   mutation        314
FT                   /*tag=  f
FT                   /note= "if this base is mutated to C, then the
FT                          protein contains the mutation I105T, which
FT                          is associated with an iron overload disorder"
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
DR   P-PSDB; AAB19149.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Disclosure; Page 2-3; 55pp; English.
XX
CC   The present sequence encodes a human histocompatibility iron loading
CC   (HFE) protein. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 2506 BP; 648 A; 552 C; 596 G; 710 T; 0 other;
```

```
  Query Match             84.4%;  Score 1215.8;  DB 21;  Length 2506;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1217;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgggcccgcgagccaggccggcgcttctcctcctgatgcttttgcagaccgcggtcctg 60

Qy    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 caggggcgcttgctgcgttcacactctctgcactacctcttcatgggtgcctcagagcag 120

Qy    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gaccttggtctttccttgtttgaagctttgggctacgtggatgaccagctgttcgtgttc 180

Qy    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tatgatcatgagagtcgccgtgtggagccccgaactccatgggtttccagtagaatttca 240

Qy    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 agccagatgtggctgcagctgagtcagagtctgaaagggtgggatcacatgttcactgtt 300

Qy    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gacttctggactattatggaaaatcacaaccacagcaaggagtcccacaccctgcaggtc 360

Qy    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641
```

```
Db    361 atcctgggctgtgaaatgcaagaagacaacagtaccgagggctactggaagtacgggtat 420

Qy    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Db    421 gatgggcaggaccaccttgaattctgccctgacacactggattggagagcagcagaaccc 480

Qy    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Db    481 agggcctggcccaccaagctggagtgggaaaggcacaagattcgggccaggcagaacagg 540

Qy    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Db    541 gcctacctggagagggactgccctgcacagctgcagcagttgctggagctgggagaggt 600

Qy    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Db    601 gttttggaccaacaagtgcctcctttggtgaaggtgacacatcatgtgacctcttcagtg 660

Qy    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Db    661 accactctacggtgtcgggccttgaactactacccccagaacatcaccatgaagtggctg 720

Qy    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Db    721 aaggataagcagccaatggatgccaaggagttcgaacctaaagacgtattgcccaatggg 780

Qy   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Db    781 gatgggacctaccagggctggataaccttggctgtacccccctggggaagagcagagatat 840

Qy   1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Db    841 acgtgccaggtggagcacccaggcctggatcagcccctcattgtgatctgggagccctca 900

Qy   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Db    901 ccgtctggcaccctagtcattggagtcatcagtggaattgctgttttgtcgtcatcttg 960

Qy   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Db    961 ttcattggaattttgttcataatattaaggaagaggcagggttcaagaggagccatgggg 1020

Qy   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Db   1021 cactacgtcttagctgaacgtgagtgacacgcagcctgcagactcactgtgggaaggaga 1080

Qy   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Db   1081 caaaactagagactcaaagagggagtgcatttatgagctcttcatgtttcaggagagagt 1140

Qy   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Db   1141 tgaacctaaacatagaaattgcctgacgaactccttgattttagccttctctgttcattt 1200

Qy   1422 CCTCAAAAAGATTTCCCCA 1440

Db   1201 cctcaaaaagatttcccca 1219


RESULT   8
AAI63897
ID   AAI63897 standard; cDNA; 596 BP.
XX
AC   AAI63897;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 105.
XX
KW   Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW   neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW   antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
KW   antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW   cardiovascular disorder; neurological disease; infection; human; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155308-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01309.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
```

PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488781/53.
DR P-PSDB; AAM43591.
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders -
XX
PS Claim 1; SEQ ID NO 105; 664pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI63803-AAI64012) and
CC the encoded proteins (AAM434497-AAM43660) useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. The
CC genes were isolated from a range of human tissues disclosed in the
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 596 BP; 133 A; 157 C; 175 G; 126 T; 5 other;

Query Match 38.7%; Score 557.6; DB 22; Length 596;
Best Local Similarity 97.6%; Pred. No. 7.4e-152;
Matches 571; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

```
Qy    166 CCCGCCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGG 225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     13 cccgccccccaaaagaagcggagatttaacggggacgtgcggccagagctggggaaatgg 72

Qy    226 GCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGG 285
          |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db     73 gcccgcgagccaggccggcgcttctcctcctgatgcttttgcakaccgcggtcctgcagg 132

Qy    286 GGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACC 345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 ggcgcttgctgcgttcacactctctgcactacctcttcatgggtgcctcagagcaggacc 192

Qy    346 TTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATG 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 ttggtctttccttgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatg 252

Qy    406 ATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCC 465
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 atcatgagagtcgccgtgtggagccccgaactccatgggtttccagtagaatttcaagcc 312

Qy    466 AGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACT 525
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    313 agatgtggctgcagctgagtcagagtctgaaa-ggtgggatcacatgttcactgttgact 371

Qy    526 TCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCC 585
          ||||||||||||||:||||||||||||||||||:||:||||||||||||||     |
Db    372 tctggactattatggraaatcacaaccacagcaaggrgtyccacaccctgcagtcatcct 431

Qy    586 TGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATG 645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 ggggctgtgaaatgcaagaagacaacagtaccgagggctactggaagtacgggtatgatg 491
```

```
Qy    646 GGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGG 705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 ggcaggaccaccttgaattctgccctgacacactggattggagagcagcagaacccaggg 551

Qy    706 CCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCA 750
          |||||||||||||| ||||||||||||||||||||||||||||||
Db    552 cctggcccaccaanctggagtgggaaaggcacaagattcgggcca 596


RESULT  9
AAC68427
ID   AAC68427 standard; DNA; 10825 BP.
XX
AC   AAC68427;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36871.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;


  Query Match             22.3%;  Score 321;  DB 22;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 2.2e-82;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtc 3881

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy    538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
          |||||||||||||||||||||||
Db   4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy    561 ---------------------------------------------------------- 560

Db   4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy    561 ---------------------------------------------------------- 560

Db   4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                             ||||||||
Db   4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy    629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          |||||||||||||||||||||||||||||
Db   4482 gctggggagaggtgttttggaccaacaag 4510


RESULT  10
AAC68428
ID   AAC68428 standard; DNA; 10825 BP.
XX
AC   AAC68428;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
```

```
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36872.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2999 A; 2252 C; 2648 G; 2926 T; 0 other;


  Query Match             22.3%;  Score 321;  DB 22;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 2.2e-82;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatgatgagagtc 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||
Db    4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy     561 ------------------------------------------------------------ 560

Db    4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy     561 ------------------------------------------------------------ 560

Db    4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy     561 ----------------------------------------------------------GAGTCCCA 568
                                                                       ||||||||
Db    4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
```

```
Db    4482 gctggggagaggtgttttggaccaacaag 4510


RESULT  11
AAT96690
ID   AAT96690 standard; DNA; 10825 BP.
XX
AC   AAT96690;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             361..7147
FT                   /*tag= a
FT                   /note= "contains introns"
FT   intron          437..3761
FT                   /*tag= b
FT                   /number= 1
FT   intron          4026..4234
FT                   /*tag= c
FT                   /number= 2
FT   intron          4511..5605
FT                   /*tag= d
FT                   /number= 3
FT   intron          5882..6039
FT                   /*tag= e
FT                   /number= 4
FT   intron          6154..7106
FT                   /*tag= f
FT                   /number= 5
FT   mutation        3872
FT                   /*tag= g
FT                   /note= "C to G substitution (24d2 mutation)
FT                         results in His to Asp substitution"
FT   variation       3878
FT                   /*tag= h
FT                   /note= "A to T substitution (24d7 variant)
FT                         results in Ser to Cys substitution"
FT   mutation        5834
FT                   /*tag= i
FT                   /note= "G to A substitution (24d1 mutation
FT                         associated with HH), results in Cys to
FT                         Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   P-PSDB; AAW36499.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 3; 115pp; English.
```

```
XX
CC   This genomic DNA sequence corresponds to the human gene whose
CC   mutated form is associated with hereditary haemochromatosis (HH).
CC   To identify this novel gene, allelic association patterns were
CC   determined between known markers and the HH locus in the HLA region
CC   of chromosome 6.  A physical clone coverage was then generated
CC   extending from D6S265, which is a marker that is centromeric of
CC   HLA-A, in a telomeric direction through D6S276, a marker at which
CC   the allelic association was no longer observed.  A single mutation
CC   (24d1) in the HH gene appears responsible for the majority of HH
CC   disease.  This comprises a G to A substitution that is present in
CC   86% of affected chromosomes and in 4% of unaffected chromosomes.
CC   It results in a Cys to Tyr substitution in the encoded protein (see
CC   AAW36499) at a critical disulphide bridge important for secondary
CC   structure.  The following are claimed: the HH genomic DNA (I), a
CC   1437 bp cDNA sequence (Ia) (see AAT96691) and their 24d1, 24d2 and
CC   24d7 variants; a cloning or expression vector; host cells; a
CC   peptide product chosen from the HH gene product, its variants
CC   (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC   residues of these; an antibody produced using the peptide; a method
CC   to determine the presence or absence of the common HH gene
CC   mutation; an animal model for the HH disease; metal chelation
CC   agents, T-cell differentiation factors and therapeutic agents for
CC   the mitigation of injury due to oxidative process in vivo or
CC   mitigation of iron overload; a method for screening potential
CC   therapeutic agents for activity in connection with HH disease; an
CC   antisense oligonucleotide directed against a transcriptional
CC   product of a nucleic acid sequence as above; and oligonucleotides
CC   or pairs of oligonucleotides covering a range of nucleotides from
CC   (I), (Ia) or their variants, useful for detecting a polymorphism in
CC   the HH gene.  The invention also relates to methods for screening
CC   for HH homozygotes, to HH diagnosis, prenatal screening and
CC   diagnosis, and therapies of HH disease, including gene therapy,
CC   protein- and antibody-based therapeutics, and small molecule
CC   therapeutics.
XX
SQ   Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;

  Query Match             22.2%;  Score 319.4;  DB 18;  Length 10825;
  Best Local Similarity   72.0%;  Pred. No. 6.4e-82;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||||
Db    4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy     561 ---------------------------------------------------------- 560

Db    4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy     561 ---------------------------------------------------------- 560

Db    4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                               ||||||||

Db    4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db    4482 gctggggagaggtgttttggaccaacaag 4510


RESULT  12
AAC68425
ID   AAC68425 standard; DNA; 10825 BP.
XX
AC   AAC68425;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36869.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;

  Query Match             22.2%;  Score 319.4;  DB 22;  Length 10825;
```

```
  Best Local Similarity   72.0%;  Pred. No. 6.4e-82;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||||
Db    4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy     561 ---------------------------------------------------------- 560

Db    4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy     561 ---------------------------------------------------------- 560

Db    4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                               ||||||||
Db    4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db    4482 gctggggagaggtgttttggaccaacaag 4510


RESULT  13
AAC68426
ID   AAC68426 standard; DNA; 10825 BP.
XX
AC   AAC68426;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36870.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;


  Query Match             22.2%;  Score 319.4;  DB 22;  Length 10825;
  Best Local Similarity   72.0%;  Pred. No. 6.4e-82;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    3822 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3942 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||||
Db    4002 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4061

Qy     561 ---------------------------------------------------------- 560

Db    4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 4121

Qy     561 ---------------------------------------------------------- 560

Db    4122 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 4181

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                               ||||||||
Db    4182 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 4361

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 4421

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 4481

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          |||||||||||||||||||||||||||||
Db   4482 gctggggagaggtgttttggaccaacaag 4510


RESULT 14
AAA96794
ID   AAA96794 standard; cDNA; 12146 BP.
XX
AC   AAA96794;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1028..1324
FT                   /*tag=  a
FT                   /number= 1
FT   intron          1325..4651
FT                   /*tag=  b
FT                   /number= 1
FT   exon            4652..4915
FT                   /*tag=  c
FT                   /number= 2
FT   intron          4916..5124
FT                   /*tag=  d
FT                   /number= 2
FT   exon            5125..5400
FT                   /*tag=  e
FT                   /number= 3
FT   intron          5401..6493
FT                   /*tag=  f
FT                   /number= 3
FT   exon            6494..6769
FT                   /*tag=  g
FT                   /number= 4
FT   intron          6770..6927
FT                   /*tag=  h
FT                   /number= 4
FT   exon            6928..7041
FT                   /*tag=  i
FT                   /number= 5
FT   intron          7042..7994
FT                   /*tag=  j
FT                   /number= 5
FT   exon            7995..9050
FT                   /*tag=  k
FT                   /number= 6
FT   intron          9051..10205
FT                   /*tag=  l
FT                   /number= 6
FT   exon            10206..10637
FT                   /*tag=  m
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Example 1; Page 21-28; 55pp; English.
XX
CC   The present sequence represents the human histocompatibility iron
CC   loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;


  Query Match             22.2%;  Score 319.4;  DB 21;  Length 12146;
  Best Local Similarity   72.0%;  Pred. No. 6.8e-82;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels  209;  Gaps    1;

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4652 gttcacactctctgcactacctcttcatgggtgcctcagagcaggaccttggtctttcct 4711

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   4712 tgtttgaagctttgggctacgtggatgaccagctgttcgtgttctatgatcatgagagtc 4771

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4772 gccgtgtggagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 4831

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4832 agctgagtcagagtctgaaagggtgggatcacatgttcactgttgacttctggactatta 4891

Qy    538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
          |||||||||||||||||||||
Db   4892 tggaaaatcacaaccacagcaagggtatgtggagaggggcctcaccttcctgaggttgt 4951

Qy    561 ---------------------------------------------------------- 560

Db   4952 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagtctgaggtcttgtg 5011

Qy    561 ---------------------------------------------------------- 560

Db   5012 ggagcagggaagagggaaggaatttgcttcctgagatcatttggtccttggggatggtgg 5071

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                             ||||||||
Db   5072 aaatagggacctattcctttggttgcagttaacaaggctggggatttttccagagtccca 5131

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5132 caccctgcaggtcatcctgggctgtgaaatgcaagaagacaacagtaccgagggctactg 5191

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5192 gaagtacgggtatgatgggcaggaccaccttgaattctgccctgacacactggattggag 5251

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5252 agcagcagaacccagggcctggcccaccaagctggagtgggaaaggcacaagattcgggc 5311

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5312 caggcagaacagggcctacctggagagggactgccctgcacagctgcagcagttgctgga 5371

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          |||||||||||||||||||||||||||||
Db    5372 gctggggagaggtgttttggaccaacaag 5400


RESULT 15
AAV57903/c
ID   AAV57903 standard; DNA; 237326 BP.
XX
AC   AAV57903;
XX
DT   21-DEC-1998  (first entry)
XX
DE   Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW   Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW   diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW   BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW   type 1 sodium transport gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO9814466-A1.
XX
PD   09-APR-1998.
XX
PF   30-SEP-1997;   97WO-US17658.
XX
PR   07-MAY-1997;   97US-0852495.
PR   01-OCT-1996;   96US-0724394.
XX
PA   (PROG-) PROGENTIOR INC.
XX
PI   Feder JN,  Kronmal GS,  Lauer PM,  Ruddy DA,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1998-240014/21.
XX
PT   Hereditary haemochromatosis gene products - used to develop products
PT   for the diagnosis and treatment of hereditary disorders in iron
PT   metabolism
XX
PS   Claim 1; Fig 9; 209pp; English.
XX
CC   The present invention describes hereditary haemochromatosis gene
CC   products from the human haemochromatosis gene. The present sequence
CC   represents a hereditary haemochromatosis subregion from an hereditary
CC   haemochromatosis (HH) affected individual. Also described is a
CC   method to determine the presence or absence of the common hereditary
CC   haemochromatosis (HFE) gene mutation in an individual comprising:
CC   (a) providing DNA or RNA from the individual; and (b) assessing the
CC   DNA or RNA for the presence or absence of a haplotype or genotype where
CC   the presence or absence of the haplotype genotype indicates the likely
CC   presence of the HFE gene mutation in the genome of the individual. The
CC   HFE gene sequences from the present invention can be used to develop
CC   products for use in the diagnosis and treatment of HFE. The present
CC   invention also describes BTF genes, which are homologues of the milk
CC   protein butyrophilin (BT), and can be used in the production of agonists
CC   and antagonists of BT function. Also described are: (1) a RoRet gene
CC   which can be used to develop products for the study, diagnosis and
CC   treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC   which are homologues of a type 1 sodium transport gene, and can
CC   similarly be used for hypophosphatemia.
XX
SQ   Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;


  Query Match           22.2%;  Score 319.4;  DB 19;  Length 237326;
  Best Local Similarity  72.0%;  Pred. No. 3e-81;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43338 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 43279

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   43278 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTC 43219

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43218 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 43159

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43158 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 43099

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------------- 560
          |||||||||||||||||||||||
Db   43098 TGGAAAATCACAACCACAGCAAGGGTATGTGGAGAGGGGGCCTCACCTTCCTGAGGTTGT 43039

Qy     561 ---------------------------------------------------------------- 560

Db   43038 CAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGGAAGTCTGAGGTCTTGTG 42979

Qy     561 ---------------------------------------------------------------- 560

Db   42978 GGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGATGGTGG 42919

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                               ||||||||
Db   42918 AAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCA 42859

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42858 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 42799

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42798 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 42739

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42738 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 42679

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42678 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 42619

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          |||||||||||||||||||||||||||||
Db   42618 GCTGGGGAGAGGTGTTTTGGACCAACAAG 42590


Search completed: June 19, 2002, 08:53:55
Job time: 8120 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:40:03 ; Search time 79.08 Seconds
(without alignments)
4472.841 Million cell updates/sec

Title: US-09-497-957-12
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1440	100.0	1440	3	US-08-652-265-12	Sequence 12, Appl
2	1440	100.0	1440	3	US-08-834-497A-12	Sequence 12, Appl
3	1440	100.0	1440	4	US-09-503-444A-12	Sequence 12, Appl
4	1438.4	99.9	1440	3	US-08-652-265-10	Sequence 10, Appl
5	1438.4	99.9	1440	3	US-08-652-265-11	Sequence 11, Appl
6	1438.4	99.9	1440	3	US-08-834-497A-10	Sequence 10, Appl
7	1438.4	99.9	1440	3	US-08-834-497A-11	Sequence 11, Appl
8	1438.4	99.9	1440	4	US-09-503-444A-10	Sequence 10, Appl
9	1438.4	99.9	1440	4	US-09-503-444A-11	Sequence 11, Appl
10	1436.8	99.8	1440	3	US-08-652-265-9	Sequence 9, Appli
11	1436.8	99.8	1440	3	US-08-834-497A-9	Sequence 9, Appli
12	1436.8	99.8	1440	4	US-09-503-444A-9	Sequence 9, Appli
13	1215.8	84.4	2506	4	US-09-277-457-1	Sequence 1, Appli
14	321	22.3	10825	3	US-08-652-265-5	Sequence 5, Appli
15	321	22.3	10825	3	US-08-652-265-7	Sequence 7, Appli
16	321	22.3	10825	3	US-08-834-497A-5	Sequence 5, Appli
17	321	22.3	10825	3	US-08-834-497A-7	Sequence 7, Appli
18	321	22.3	10825	4	US-09-503-444A-5	Sequence 5, Appli
19	321	22.3	10825	4	US-09-503-444A-7	Sequence 7, Appli
20	319.4	22.2	10825	3	US-08-652-265-1	Sequence 1, Appli
21	319.4	22.2	10825	3	US-08-652-265-3	Sequence 3, Appli
22	319.4	22.2	10825	3	US-08-834-497A-1	Sequence 1, Appli
23	319.4	22.2	10825	3	US-08-834-497A-3	Sequence 3, Appli
24	319.4	22.2	10825	4	US-09-503-444A-1	Sequence 1, Appli
25	319.4	22.2	10825	4	US-09-503-444A-3	Sequence 3, Appli
26	319.4	22.2	12146	4	US-09-277-457-27	Sequence 27, Appl
27	319.4	22.2	246240	2	US-08-724-394A-20	Sequence 20, Appl
28	319.4	22.2	246240	2	US-08-724-394A-21	Sequence 21, Appl
29	319.4	22.2	246240	2	US-08-724-394A-22	Sequence 22, Appl
30	280	19.4	517	1	US-08-632-673B-4	Sequence 4, Appli
31	280	19.4	517	3	US-08-652-265-21	Sequence 21, Appl
32	280	19.4	517	3	US-08-834-497A-21	Sequence 21, Appl
33	280	19.4	517	4	US-09-503-444A-21	Sequence 21, Appl
34	279	19.4	517	1	US-08-632-673B-13	Sequence 13, Appl
35	278.4	19.3	360	3	US-08-905-124-5	Sequence 5, Appli
36	278.4	19.3	517	1	US-08-632-673B-3	Sequence 3, Appli
37	278.4	19.3	517	3	US-08-652-265-20	Sequence 20, Appl
38	278.4	19.3	517	3	US-08-834-497A-20	Sequence 20, Appl
39	278.4	19.3	517	4	US-09-503-444A-20	Sequence 20, Appl
40	173.6	12.1	1112	3	US-08-890-719-5	Sequence 5, Appli
41	172	11.9	1230	3	US-08-890-719-9	Sequence 9, Appli
42	168.8	11.7	1195	3	US-08-890-719-10	Sequence 10, Appl
43	166.8	11.6	1197	3	US-08-890-719-7	Sequence 7, Appli
44	166.8	11.6	1262	3	US-08-890-719-37	Sequence 37, Appl
45	162	11.2	1145	3	US-08-890-719-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT   1
US-08-652-265-12
; Sequence 12, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
```

```
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-12


  Query Match            100.0%;  Score 1440;  DB 3;  Length 1440;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   2
US-08-834-497A-12
; Sequence 12, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
```

```
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-12


  Query Match             100.0%;  Score 1440;  DB 3;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440

RESULT   3
US-09-503-444A-12
; Sequence 12, Application US/09503444A
```

```
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-12


  Query Match           100.0%;  Score 1440;  DB 4;  Length 1440;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   4
US-08-652-265-10
; Sequence 10, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-10


  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   5
US-08-652-265-11
; Sequence 11, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-652-265-11


  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
```

```
Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT   6
US-08-834-497A-10
; Sequence 10, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-10
```

```
  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420
```

```
Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT  7
US-08-834-497A-11
; Sequence 11, Application US/08834497A
```

```
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-834-497A-11


  Query Match             99.9%;  Score 1438.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
```

```
Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT   8
US-09-503-444A-10
; Sequence 10, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-10


  Query Match             99.9%;  Score 1438.4;  DB 4;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT   9
US-09-503-444A-11
; Sequence 11, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  08/652,265
;       FILING DATE:  23-May-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/632,673
;       FILING DATE:  16-Apr-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/630,912
;       FILING DATE:  04-Apr-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Poissant, Brian M.
;       REGISTRATION NUMBER:  28,462
;       REFERENCE/DOCKET NUMBER:  8907-0088-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-790-9090
;       TELEFAX:  212-869-9741
;       TELEX:  66141
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1440 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  222..1268
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(408, "g")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d2
US-09-503-444A-11


  Query Match             99.9%;  Score 1438.4;  DB 4;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1439;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420

Qy     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440


RESULT  10
US-08-652-265-9
; Sequence 9, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
```

```
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-9


  Query Match             99.8%;  Score 1436.8;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1438;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
```

```
Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 11
US-08-834-497A-9
; Sequence 9, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-9
```

```
  Query Match             99.8%;  Score 1436.8;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1438;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240

Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
```

```
Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320

Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT  12
US-09-503-444A-9
; Sequence 9, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-9
```

```
  Query Match             99.8%;  Score 1436.8;  DB 4;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1438;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGGACACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60

Qy     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAG 120

Qy    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAG 180

Qy    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGCGAGCCAGGC 240
```

```
Qy    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTGCTGCGTT 300

Qy    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360

Qy    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCC 420
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC 420

Qy    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGC 600

Qy    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTG 660

Qy    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGC 720

Qy    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780

Qy    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Qy    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGG 900

Qy    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960

Qy    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCT 1020

Qy   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC 1080
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   1021 GGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCA 1200

Qy   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAAC 1260

Qy   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAG 1320
```

```
Qy   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCA 1440
```

RESULT 13
US-09-277-457-1
; Sequence 1, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
; APPLICANT: Rothenberg, Barry E.
; APPLICANT: Sawada-Hirai, Ritsuko
; APPLICANT: Barton, James C.
; TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
; FILE REFERENCE: 10653/002001
; CURRENT APPLICATION NUMBER: US/09/277,457
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2506
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: Missense mutation at nucleotide 314
US-09-277-457-1

Query Match 84.4%; Score 1215.8; DB 4; Length 2506;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1217; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgggcccgcgagccaggccggcgcttctcctcctgatgcttttgcagaccgcggtcctg 60

Qy    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 caggggcgcttgctgcgttcacactctctgcactacctcttcatgggtgcctcagagcag 120

Qy    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gaccttggtctttccttgtttgaagctttgggctacgtggatgaccagctgttcgtgttc 180

Qy    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tatgatcatgagagtcgccgtgtggagccccgaactccatgggtttccagtagaatttca 240

Qy    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 agccagatgtggctgcagctgagtcagagtctgaaagggtgggatcacatgttcactgtt 300

Qy    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gacttctggactattatggaaaatcacaaccacagcaaggagtcccacaccctgcaggtc 360

Qy    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 atcctgggctgtgaaatgcaagaagacaacagtaccgagggctactggaagtacgggtat 420

Qy    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gatgggcaggaccaccttgaattctgccctgacacactggattggagagcagcagaaccc 480

Qy    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 agggcctggcccaccaagctggagtgggaaaggcacaagattcgggccaggcagaacagg 540
```

```
Qy     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 gcctacctggagagggactgccctgcacagctgcagcagttgctggagctggggagaggt 600

Qy     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gttttggaccaacaagtgcctcctttggtgaaggtgacacatcatgtgacctcttcagtg 660

Qy     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 accactctacggtgtcgggccttgaactactacccccagaacatcaccatgaagtggctg 720

Qy     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 aaggataagcagccaatggatgccaaggagttcgaacctaaagacgtattgcccaatggg 780

Qy    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gatgggacctaccagggctggataaccttggctgtacccccctggggaagagcagagatat 840

Qy    1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 acgtgccaggtggagcacccaggcctggatcagccccctcattgtgatctgggagccctca 900

Qy    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ccgtctggcaccctagtcattggagtcatcagtggaattgctgtttttgtcgtcatcttg 960

Qy    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ttcattggaattttgttcataatattaaggaagaggcagggttcaagaggagccatgggg 1020

Qy    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 cactacgtcttagctgaacgtgagtgacacgcagcctgcagactcactgtgggaaggaga 1080

Qy    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 caaaactagagactcaaagagggagtgcatttatgagctcttcatgtttcaggagagagt 1140

Qy    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 tgaacctaaacatagaaattgcctgacgaactccttgattttagccttctctgttcattt 1200

Qy    1422 CCTCAAAAAGATTTCCCCA 1440
           |||||||||||||||||||
Db    1201 cctcaaaaagatttcccca 1219
```

```
RESULT  14
US-08-652-265-5
; Sequence 5, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;    • ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-652-265-5
```

```
  Query Match             22.3%;  Score 321;  DB 3;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 4.8e-83;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy     298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3762 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 3821

Qy     358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3822 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 3881

Qy     418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3882 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 3941

Qy     478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3942 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 4001

Qy    538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
          |||||||||||||||||||||||
Db   4002 TGGAAAATCACAACCACAGCAAGGGTATGTGGAGAGGGGGCCTCACCTTCCTGAGGTTGT 4061

Qy    561 ------------------------------------------------------------ 560

Db   4062 CAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGGAAGTCTGAGGTCTTGTG 4121

Qy    561 ------------------------------------------------------------ 560

Db   4122 GGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGATGGTGG 4181

Qy    561 ----------------------------------------------------GAGTCCCA 568
                                                                 ||||||||
Db   4182 AAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCA 4241

Qy    569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4242 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 4301

Qy    629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 4361

Qy    689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 4421

Qy    749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 4481

Qy    809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
          |||||||||||||||||||||||||||||
Db   4482 GCTGGGGAGAGGTGTTTTGGACCAACAAG 4510


RESULT  15
US-08-652-265-7
; Sequence 7, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-7


  Query Match             22.3%;  Score 321;  DB 3;  Length 10825;
  Best Local Similarity   72.1%;  Pred. No. 4.8e-83;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels  209;  Gaps    1;

Qy    298 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3762 GTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCT 3821

Qy    358 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3822 TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTC 3881

Qy    418 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3882 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 3941

Qy    478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    3942 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 4001

Qy     538 TGGAAAATCACAACCACAGCAAG---------------------------------- 560
           |||||||||||||||||||||||
Db    4002 TGGAAAATCACAACCACAGCAAGGGTATGTGGAGAGGGGGCCTCACCTTCCTGAGGTTGT 4061

Qy     561 ------------------------------------------------------------ 560

Db    4062 CAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGGAAGTCTGAGGTCTTGTG 4121

Qy     561 ------------------------------------------------------------ 560

Db    4122 GGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGATGGTGG 4181

Qy     561 ----------------------------------------------------GAGTCCCA 568
                                                                  ||||||||
Db    4182 AAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCA 4241

Qy     569 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4242 CACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 4301

Qy     629 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4302 GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG 4361

Qy     689 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4362 AGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 4421

Qy     749 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4422 CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 4481

Qy     809 GCTGGGGAGAGGTGTTTTGGACCAACAAG 837
           |||||||||||||||||||||||||||||
Db    4482 GCTGGGGAGAGGTGTTTTGGACCAACAAG 4510
```

Search completed: June 19, 2002, 08:40:57
Job time: 7342 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:40:02 ; Search time 1828.83 Seconds
(without alignments)
3444.217 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870
Perfect score: 301
Sequence: 1 GGCACGGAATCCCTGGTTGG.....GCACTTTGGGAGGCCGAGGC 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	301	100.0	10825	6	AR117789	AR117789 Sequence
2	301	100.0	10825	6	AR117790	AR117790 Sequence
3	301	100.0	10825	6	AR117791	AR117791 Sequence
4	301	100.0	10825	6	AR117792	AR117792 Sequence
5	301	100.0	10825	6	AR149459	AR149459 Sequence
6	301	100.0	10825	6	AR149460	AR149460 Sequence
7	301	100.0	10825	6	AR149461	AR149461 Sequence
8	301	100.0	10825	6	AR149462	AR149462 Sequence
9	301	100.0	12146	9	HSHFE	Z92910 Homo sapien
10	301	100.0	246240	6	AR036572	AR036572 Sequence
11	301	100.0	246240	6	AR036573	AR036573 Sequence
12	301	100.0	246240	6	AR036574	AR036574 Sequence
c 13	301	100.0	246282	9	HSU91328	U91328 Human hered
c 14	299.4	99.5	193752	2	AL359892	AL359892 Homo sapi
15	88.8	29.5	4349	4	AF301592	AF301592 Diceros b
16	70.8	23.5	181168	2	AC055888	AC055888 Homo sapi
c 17	70.8	23.5	186731	2	AC073212	AC073212 Homo sapi
c 18	70.8	23.5	187818	2	AC093333	AC093333 Homo sapi
19	68.6	22.8	164155	9	AC000385	AC000385 Homo sapi
c 20	68.4	22.7	182211	9	AL360004	AL360004 Human DNA
21	68.2	22.7	176379	2	AC060833	AC060833 Homo sapi
c 22	68.2	22.7	183239	2	AC015906	AC015906 Homo sapi
c 23	67.6	22.5	174006	2	AC027062	AC027062 Homo sapi
c 24	67.6	22.5	182288	9	AC018362	AC018362 Homo sapi
25	67.2	22.3	120871	9	AC016605	AC016605 Homo sapi
26	67.2	22.3	231464	9	AF111168	AF111168 Homo sapi
c 27	66.6	22.1	74517	9	HSAJ9615	AJ009615 Homo sapi
c 28	66.6	22.1	134026	2	AC069317	AC069317 Homo sapi
c 29	66.6	22.1	155862	9	AC018464	AC018464 Homo sapi
30	66.6	22.1	170616	2	AC015922	AC015922 Homo sapi
c 31	66.6	22.1	185415	9	HS121M24	AL354046 Homo sapi
32	66.4	22.1	143811	2	AC013617	AC013617 Homo sapi
33	66.4	22.1	154443	9	AC087315	AC087315 Homo sapi
c 34	66.4	22.1	154848	9	AC016555	AC016555 Homo sapi
c 35	66.4	22.1	183624	2	AC084116	AC084116 Homo sapi
36	66.4	22.1	191130	2	AC010897	AC010897 Homo sapi
c 37	66.4	22.1	195668	9	AL353729	AL353729 Human DNA
c 38	66.4	22.1	200673	9	AC021396	AC021396 Homo sapi
c 39	66.2	22.0	148456	2	HS378I6	AL591222 Homo sapi
40	66	21.9	193963	9	AC013475	AC013475 Homo sapi
c 41	65.6	21.8	93847	9	AL157373	AL157373 Human DNA
c 42	65.6	21.8	123050	9	AC008917	AC008917 Homo sapi
43	65.6	21.8	150409	9	AP003496	AP003496 Homo sapi
c 44	65.6	21.8	175582	9	AC023590	AC023590 Homo sapi
45	65.6	21.8	201300	9	AL137073	AL137073 Human DNA

ALIGNMENTS

RESULT 1
AR117789
LOCUS AR117789 10825 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6140305.
ACCESSION AR117789
VERSION AR117789.1 GI:14098695
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 1 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. .10825
/organism="unknown"
BASE COUNT 2998 a 2253 c 2648 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;

```
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
           |
Db    4870 C 4870


RESULT   2
AR117790
LOCUS       AR117790               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 3 from patent US 6140305.
ACCESSION   AR117790
VERSION     AR117790.1  GI:14098696
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 3 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2253 c   2647 g   2926 t
ORIGIN


  Query Match            100.0%;  Score 301;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
           |
Db    4870 C 4870


RESULT   3
AR117791
LOCUS       AR117791               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 5 from patent US 6140305.
ACCESSION   AR117791
VERSION     AR117791.1  GI:14098697
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 5 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2252 c   2649 g   2926 t
ORIGIN


  Query Match            100.0%;  Score 301;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
           |
Db    4870 C 4870


RESULT   4
AR117792
LOCUS       AR117792               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 7 from patent US 6140305.
ACCESSION   AR117792
VERSION     AR117792.1  GI:14098698
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 7 31-OCT-2000;
FEATURES             Location/Qualifiers
```

```
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2252 c   2648 g   2926 t
ORIGIN


  Query Match             100.0%;  Score 301;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
          |
Db    4870 C 4870


RESULT   5
AR149459
LOCUS       AR149459               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 1 from patent US 6228594.
ACCESSION   AR149459
VERSION     AR149459.1  GI:15114050
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 1 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2253 c   2648 g   2926 t
ORIGIN


  Query Match             100.0%;  Score 301;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
          |
Db    4870 C 4870


RESULT   6
AR149460
LOCUS       AR149460               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 3 from patent US 6228594.
ACCESSION   AR149460
VERSION     AR149460.1  GI:15114051
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 3 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2253 c   2647 g   2926 t
ORIGIN


  Query Match             100.0%;  Score 301;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
          |
Db    4870 C 4870


RESULT   7
AR149461
LOCUS       AR149461               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 5 from patent US 6228594.
ACCESSION   AR149461
VERSION     AR149461.1  GI:15114052
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 5 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2252 c   2649 g   2926 t
ORIGIN
```

```
  Query Match             100.0%;  Score 301;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870
```

```
RESULT   8
AR149462
LOCUS       AR149462              10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 7 from patent US 6228594.
ACCESSION   AR149462
VERSION     AR149462.1  GI:15114053
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 7 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2252 c   2648 g   2926 t
ORIGIN
```

```
  Query Match             100.0%;  Score 301;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870
```

```
RESULT   9
HSHFE
LOCUS       HSHFE                 12146 bp    DNA     linear   PRI 23-JUL-1999
DEFINITION  Homo sapiens HFE gene.
ACCESSION   Z92910
VERSION     Z92910.1  GI:1890179
KEYWORDS    haemochromatosis; HFE gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 858)
  AUTHORS   Albig,W., Drabent,B., Burmester,N., Bode,C. and Doenecke,D.
  TITLE     The haemochromatosis candidate gene HFE (HLA-H) of man and mouse is
            located in syntenic regions within the histone gene cluster
  JOURNAL   J. Cell. Biochem. 69 (2), 117-126 (1998)
  MEDLINE   98208340
REFERENCE   2  (bases 1 to 12146)
  AUTHORS   Albig,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-1997) Albig W., Georg-August-Universitaet
            Goettingen, Biochemie und Molekulare Zellbiologie, Humboldtallee
            23, Goettingen, FRG, 37073
FEATURES             Location/Qualifiers
     source          1. .12146
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p"
                     /clone="ICRFy901D1223"
                     /clone_lib="ICRF YAC-library"
     exon            1028. .1324
                     /gene="HFE"
                     /number=1
     gene            1028. .10637
                     /gene="HFE"
     CDS             join(1249. .1324,4652. .4915,5125. .5400,6494. .6769,
                     6928. .7041,7995. .8035)
                     /gene="HFE"
                     /function="iron metabolism"
                     /note="haemochromatosis candidate gene"
                     /codon_start=1
                     /protein_id="CAB07442.1"
                     /db_xref="GI:1890180"
                     /db_xref="SWISS-PROT:Q30201"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
```

```
                     SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                     EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                     SRGAMGHYVLAERE"
     intron          1325. .4651
                     /gene="HFE"
                     /number=1
     repeat_unit     3494. .3735
                     /gene="HFE"
                     /rpt_family="Alu"
     repeat_unit     3973. .4283
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            4652. .4915
                     /gene="HFE"
                     /number=2
     intron          4916. .5124
                     /gene="HFE"
                     /number=2
     exon            5125. .5400
                     /gene="HFE"
                     /number=3
     intron          5401. .6493
                     /gene="HFE"
                     /number=3
     repeat_unit     5707. .6005
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            6494. .6769
                     /gene="HFE"
                     /number=4
     intron          6770. .6927
                     /gene="HFE"
                     /number=4
     exon            6928. .7041
                     /gene="HFE"
                     /number=5
     intron          7042. .7994
                     /gene="HFE"
                     /number=5
     exon            7995. .9050
                     /gene="HFE"
                     /number=6
     repeat_unit     9017. .9340
                     /gene="HFE"
                     /rpt_family="Alu"
     intron          9051. .10205
                     /gene="HFE"
                     /number=6
     repeat_unit     9957. .10239
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            10206. .10637
                     /gene="HFE"
                     /number=7
     polyA_signal    10617. .10622
                     /gene="HFE"
BASE COUNT     3383 a   2474 c   2911 g   3378 t
ORIGIN


  Query Match             100.0%;  Score 301;  DB 9;  Length 12146;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5460 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 5519

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5520 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 5579

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5580 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 5639

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5640 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 5699

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5700 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 5759

Qy     301 C 301
           |
Db    5760 C 5760


RESULT  10
AR036572
LOCUS       AR036572              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 20 from patent US 5872237.
ACCESSION   AR036572
VERSION     AR036572.1  GI:5953240
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 20 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN


  Query Match             100.0%;  Score 301;  DB 6;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196873 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 196932

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196933 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 196992

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196993 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 197052

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197053 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 197112

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197113 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 197172

Qy     301 C 301
           |
Db 197173 C 197173


RESULT  11
AR036573
LOCUS       AR036573              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 21 from patent US 5872237.
ACCESSION   AR036573
VERSION     AR036573.1  GI:5953241
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 21 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN


  Query Match             100.0%;  Score 301;  DB 6;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196873 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 196932

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196933 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 196992

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196993 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 197052

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197053 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 197112

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197113 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 197172

Qy    301 C 301
          |
Db 197173 C 197173


RESULT  12
AR036574
LOCUS       AR036574              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 22 from patent US 5872237.
ACCESSION   AR036574
VERSION     AR036574.1  GI:5953242
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 22 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN


  Query Match             100.0%;  Score 301;  DB 6;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 196873 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 196932

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196933 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 196992

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196993 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 197052

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197053 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 197112

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197113 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 197172

Qy    301 C 301
          |
Db 197173 C 197173


RESULT  13
HSU91328/c
LOCUS       HSU91328              246282 bp    DNA     linear   PRI 17-DEC-2001
DEFINITION  Human hereditary haemochromatosis region, histone 2A-like protein
            gene, hereditary haemochromatosis (HLA-H) gene, RoRet gene, and
            sodium phosphate transporter (NPT3) gene, complete cds.
ACCESSION   U91328
VERSION     U91328.1  GI:2088550
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 246282)
  AUTHORS   Ruddy,D.A., Kronmal,G.S., Lee,V.K., Mintier,G.A., Quintana,L.,
            Domingo,R. Jr., Meyer,N.C., Irrinki,A., McClelland,E.E., Fullan,A.,
            Mapa,F.A., Moore,T., Thomas,W., Loeb,D.B., Harmon,C.,
            Tsuchihashi,Z., Wolff,R.K., Schatzman,R.C. and Feder,J.N.
  TITLE     A 1.1-Mb transcript map of the hereditary hemochromatosis locus
  JOURNAL   Genome Res. 7 (5), 441-456 (1997)
  MEDLINE   97294057
   PUBMED   9149941
REFERENCE   2  (bases 1 to 246282)
  AUTHORS   Ruddy,D.A., Kronmal,G.S., Lee,V.K., Mintier,G.A., Quintana,L.,
            Domingo,R. Jr., Meyer,N.C., Irrinke,A., McClelland,E., Fullan,A.,
            Mapa,F.A., Moore,T., Thomas,W., Loeb,D.B., Harmon,C.,
            Tsuchihashi,Z., Wolff,R.K., Schatzman,R.C. and Feder,J.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1997) Sequencing, Mercator Genetics, 4040
            Campbell Avenue, Menlo Park, CA 94025, USA
FEATURES             Location/Qualifiers
     source          1. .246282
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="pac222k22; pac75I14; bac132a2"
     repeat_region   complement(1231. .1501)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   3528. .3676
                     /rpt_family="MER3"
                     /rpt_type=dispersed
     repeat_region   4056. .4304
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(5496. .5780)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   6032. .6461
                     /rpt_family="Alu"
                     /rpt_type=dispersed
```

```
repeat_region   8203. .8463
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   8507. .8616
                /rpt_family="MSTAR"
                /rpt_type=dispersed
repeat_region   complement(8888. .9194)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(9523. .9792)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   10109. .10375
                /rpt_family="MSTAR"
                /rpt_type=dispersed
repeat_region   complement(10384. .10680)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   11482. .11742
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(12899. .13326)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   15021. .15309
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(15450. .15714)
                /rpt_family="Alu"
                /rpt_type=dispersed
CDS             complement(16434. .16826)
                /codon_start=1
                /product="histone 2A-like protein"
                /protein_id="AAB82086.1"
                /db_xref="GI:2088554"
                /translation="MSGRGKQGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGNYAERV
                GAGAPVYLAAVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGRVT
                IAQGGVLPNIQAVLLPKKTESHHKAKGK"
repeat_region   complement(18988. .19282)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(20597. .20869)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(24343. .24624)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   25654. .25805
                /rpt_family="MIR"
                /rpt_type=dispersed
repeat_region   29842. .29938
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(30592. .30869)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   31223. .31507
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(31819. .32011)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   33810. .34150
                /rpt_family="MER1"
                /rpt_type=dispersed
repeat_region   34235. .34314
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(35768. .36062)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(36267. .36548)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   39132. .39252
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(40045. .40645)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   40960. .41246
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   41581. .42174
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   42432. .42541
                /rpt_family="MIR"
                /rpt_type=dispersed
repeat_region   44636. .44911
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(45532. .45851)
                /rpt_family="Alu"
                /rpt_type=dispersed
gene            complement(46833. .53618)
                /gene="HLA-H"
                /note="HFE"
CDS             complement(join(46833. .46873,47827. .47940,48099. .48374,
                49470. .49745,49955. .50218,53543. .53618))
                /gene="HLA-H"
                /codon_start=1
                /product="hereditary hemochromatosis"
                /protein_id="AAB82083.1"
                /db_xref="GI:2088551"
                /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                SRGAMGHYVLAERE"
repeat_region   complement(48868. .49182)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(50599. .50888)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(51145. .51372)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   51820. .51934
                /rpt_family="MER20"
                /rpt_type=dispersed
repeat_region   55276. .55719
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(56085. .56461)
                /rpt_family="L1"
                /rpt_type=dispersed
repeat_region   complement(56521. .56801)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   57660. .57815
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   58118. .58483
                /rpt_family="MLT1"
                /rpt_type=dispersed
repeat_region   58672. .58900
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(59299. .59737)
                /rpt_family="LTR2"
                /rpt_type=dispersed
repeat_region   complement(65665. .66013)
```

```
                     /rpt_family="LTR2"
                     /rpt_type=dispersed
     repeat_region   66764. .67053
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   67743. .68016

  Query Match             100.0%;  Score 301;  DB 9;  Length 246282;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-76;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   49410 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 49351

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   49350 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 49291

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   49290 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 49231

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   49230 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 49171

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   49170 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 49111

Qy     301 C 301
          |
Db   49110 C 49110


RESULT  14
AL359892/c
LOCUS       AL359892              193752 bp    DNA     linear   HTG 13-JUN-2001
DEFINITION  Homo sapiens chromosome 6 clone RP11-557F22, *** SEQUENCING IN
            PROGRESS ***, 18 unordered pieces.
ACCESSION   AL359892
VERSION     AL359892.5  GI:9930971
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 193752)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 27, 2000 this sequence version replaced gi:9864230.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA557F22
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 183925 bases at least Q40
            Consensus quality: 187703 bases at least Q30
            Consensus quality: 189658 bases at least Q20
            Insert size: 192052; sum-of-contigs
            Insert size: 198247; agarose-fp
            Quality coverage: 3.68x in Q20 bases; sum-of-contigs Quality
            coverage: 3.70x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3250: contig of 3250 bp in length
            *     3251 3350: gap of      100 bp
            *     3351    14600: contig of 11250 bp in length
            *    14601 14700: gap of      100 bp
            *    14701    32357: contig of 17657 bp in length
            *    32358 32457: gap of      100 bp
            *    32458    34886: contig of 2429 bp in length
            *    34887 34986: gap of      100 bp
            *    34987    43490: contig of 8504 bp in length
            *    43491 43590: gap of      100 bp
            *    43591    47437: contig of 3847 bp in length
            *    47438 47537: gap of      100 bp
            *    47538    57356: contig of 9819 bp in length
            *    57357 57456: gap of      100 bp
            *    57457    59845: contig of 2389 bp in length
            *    59846 59945: gap of      100 bp
            *    59946    63972: contig of 4027 bp in length
            *    63973 64072: gap of      100 bp
            *    64073    82711: contig of 18639 bp in length
            *    82712 82811: gap of      100 bp
            *    82812   111814: contig of 29003 bp in length
            *   111815 111914: gap of      100 bp
            *   111915   120276: contig of 8362 bp in length
            *   120277 120376: gap of      100 bp
            *   120377   136660: contig of 16284 bp in length
            *   136661 136760: gap of      100 bp
            *   136761   153913: contig of 17153 bp in length
            *   153914 154013: gap of      100 bp
            *   154014   158659: contig of 4646 bp in length
            *   158660 158759: gap of      100 bp
            *   158760   164235: contig of 5476 bp in length
            *   164236 164335: gap of      100 bp
            *   164336   184996: contig of 20661 bp in length
            *   184997 185096: gap of      100 bp
            *   185097   193752: contig of 8656 bp in length.
FEATURES             Location/Qualifiers
     source          1. .193752
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-557F22"
                     /clone_lib="RPCI-11.2"
     misc_feature    1. .3250
                     /note="assembly_fragment:01752
                     fragment_chain:1"
     misc_feature    3351. .14600
                     /note="assembly_fragment:01177
                     fragment_chain:1"
     misc_feature    14701. .32357
                     /note="assembly_fragment:00673
                     fragment_chain:1"
     misc_feature    32458. .34886
                     /note="assembly_fragment:00884
                     fragment_chain:1"
     misc_feature    34987. .43490
                     /note="assembly_fragment:00652
                     fragment_chain:1"
     misc_feature    43591. .47437
                     /note="assembly_fragment:00628
                     fragment_chain:1"
     misc_feature    47538. .57356
                     /note="assembly_fragment:00983
                     fragment_chain:2"
     misc_feature    57457. .59845
```

```
                     /note="assembly_fragment:01129
                     fragment_chain:2"
     misc_feature    59946. .63972
                     /note="assembly_fragment:01207
                     fragment_chain:3"
     misc_feature    64073. .82711
                     /note="assembly_fragment:00301
                     fragment_chain:3"
     misc_feature    82812. .111814
                     /note="assembly_fragment:00470"
     misc_feature    111915. .120276
                     /note="assembly_fragment:00783.0"
     misc_feature    120377. .136660
                     /note="assembly_fragment:01014"
     misc_feature    136761. .153913
                     /note="assembly_fragment:01161"
     misc_feature    154014. .158659
                     /note="assembly_fragment:01329"
     misc_feature    158760. .164235
                     /note="assembly_fragment:01675"
     misc_feature    164336. .184996
                     /note="assembly_fragment:01884"
     misc_feature    185097. .193752
                     /note="assembly_fragment:01893"
BASE COUNT    55929 a  40409 c  39996 g  55686 t   1732 others
ORIGIN


  Query Match             99.5%;  Score 299.4;  DB 2;  Length 193752;
  Best Local Similarity   99.7%;  Pred. No. 4.1e-76;
  Matches  300;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12082 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 12023

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12022 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 11963

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11962 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 11903

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11902 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 11843

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   11842 AATGTCAAGGCCGGGCACGGTGGCTCACCCTTGTAATCCCAGCACTTTGGGAGGCCGAGG 11783

Qy     301 C 301
          |
Db   11782 C 11782


RESULT  15
AF301592
LOCUS       AF301592               4349 bp    DNA     linear   MAM 13-DEC-2000
DEFINITION  Diceros bicornis HFE gene, exons 2 through 6 and partial cds.
ACCESSION   AF301592
VERSION     AF301592.1  GI:11692702
KEYWORDS    .
SOURCE      black rhinoceros.
  ORGANISM  Diceros bicornis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
REFERENCE   1  (bases 1 to 4349)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4349)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .4349
                     /organism="Diceros bicornis"
                     /db_xref="taxon:9805"
     mRNA            join(<1. .264,468. .743,1588. .1863,1995. .2108,3109. .>3149)
                     /product="HFE"
     exon            1. .264
                     /number=2
     CDS             join(<1. .264,468. .743,1588. .1863,1995. .2108,3109. .3149)
                     /codon_start=3
                     /product="HFE"
                     /protein_id="AAG39940.1"
                     /db_xref="GI:11692703"
                     /translation="SHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAES
                     RAQWVLGEAHSQLWLQLTQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQE
                     DNSTRGFWKYGYDGQDHLEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERD
                     CPEQLQWLLELGRGVLDQQVPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRK
                     PVDVKDAESKDVLPSGDGTYQSWEALAVPPGEEQRYTCQVEHPGLDQPLTATWEPSLS
                     NTLVTGVISGIAVCVIIFFIGILFRILRKRQASRGAMGDYVLAECE"
     exon            468. .743
                     /number=3
     exon            1588. .1863
                     /number=4
     exon            1995. .2108
                     /number=5
     exon            3109. .>3149
                     /number=6
BASE COUNT     1142 a    924 c   1124 g   1159 t
ORIGIN


  Query Match             29.5%;  Score 88.8;  DB 4;  Length 4349;
  Best Local Similarity   75.9%;  Pred. No. 2e-15;
  Matches  123;  Conservative    0;  Mismatches   37;  Indels    2;  Gaps    1;

Qy      87 GTCTCTACCTTATAATTGAGATGTATGAGACAGCCACAAGTCATGGGTTTAATTTCTTTT 146
           ||||| ||||| | |||||||||||||| ||| |||  |||||||| |||||||| ||||
Db     908 GTCTCCACCTTGTCATTGAGATGTATGACACAACCATGAGTCATGGTTTTAATTTTTTTT 967

Qy     147 CTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCCCTTGCTTTTTATTTAACCAATAA 206
           | |||  |||||||  ||   |||||||| ||  |  | || ||  |||||||| ||
Db     968 CCCCACACATATGGCCAAAGAAGAAGTGTCTGTGATC--TTCTCTTGGTTTAACCAAGAA 1025

Qy     207 TCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAA 248
           | ||| || |||||| | | | |||||||||  ||||||
Db    1026 CTCTGTGTCTACTTATACATTTCAGAAATTCAGAGGTGTCAA 1067


Search completed: June 19, 2002, 10:58:56
Job time: 8334 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:24:43 ; Search time 1700.16 Seconds
(without alignments)
2389.530 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870
Perfect score: 301
Sequence: 1 GGCACGGAATCCCTGGTTGG.....GCACTTTGGGAGGCCGAGGC 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	260.2	86.4	531	12	AQ703007	AQ703007 HS_5443_B
	2	65.2	21.7	630	12	AG019823	AG019823 Homo sapi
	3	65.2	21.7	659	12	AG020621	AG020621 Homo sapi
	4	65	21.6	422	10	W44873	W44873 zc05d01.r1
c	5	64	21.3	651	10	BG388094	BG388094 602413054
	6	63.6	21.1	533	12	AQ471624	AQ471624 CITBI-E1-
c	7	63.4	21.1	303	12	AZ756856	AZ756856 ew02e08.x
c	8	63	20.9	389	9	AA825901	AA825901 od52h11.s
c	9	62.8	20.9	342	9	AI092525	AI092525 qa38f07.x
c	10	62.8	20.9	439	12	AQ150673	AQ150673 HS_3203_A
	11	62.8	20.9	553	9	AW020486	AW020486 df10e08.y
	12	62.2	20.7	281	9	AI565873	AI565873 tq49g04.x
	13	62.2	20.7	310	10	BF898122	BF898122 RC1-MT016
	14	62.2	20.7	400	10	R89296	R89296 yq02f11.s1
	15	62.2	20.7	408	10	H71349	H71349 yu55h06.s1
	16	62.2	20.7	414	10	N58729	N58729 yv74a10.s1
c	17	62.2	20.7	556	12	AQ489145	AQ489145 RPCI-11-2
	18	62.2	20.7	603	9	AF063563	AF063563 AF063563
	19	62.2	20.7	603	9	AI110688	AI110688 HA0123 Hu
c	20	62.2	20.7	662	12	AG061477	AG061477 Pan trogl
	21	62.2	20.7	1825	11	AF113008	AF113008 Homo sapi
c	22	61.4	20.4	619	9	BE147774	BE147774 RC1-HT022
	23	61.2	20.3	466	12	AQ269092	AQ269092 RPCI11-74
c	24	61.2	20.3	703	12	AG090274	AG090274 Pan trogl
	25	61	20.3	645	12	AG095800	AG095800 Pan trogl
	26	60.8	20.2	369	10	F25183	F25183 HSPD12013 H
	27	60.6	20.1	435	9	AA573540	AA573540 nf41b05.s
	28	60.6	20.1	455	9	AI732329	AI732329 nf41b05.x
	29	60.6	20.1	511	9	AI732239	AI732239 nf35f10.x
	30	60.6	20.1	693	12	AG144127	AG144127 Pan trogl
	31	60.4	20.1	656	12	AG115691	AG115691 Pan trogl
	32	60.4	20.1	657	12	AG084435	AG084435 Pan trogl
	33	60.2	20.0	363	12	AQ385523	AQ385523 RPCI11-13
	34	60.2	20.0	586	9	AV713396	AV713396 AV713396
	35	60	19.9	368	9	AA720754	AA720754 nv93g11.r
c	36	60	19.9	482	12	AQ631584	AQ631584 RPCI-11-4
c	37	60	19.9	748	10	BI522662	BI522662 603175686
c	38	60	19.9	822	10	BI906585	BI906585 603064226
c	39	60	19.9	1110	10	BG257716	BG257716 602377406
c	40	60	19.9	1378	11	BC008448	BC008448 Homo sapi
c	41	59.8	19.9	495	10	BG054544	BG054544 7o45h12.x
c	42	59.6	19.8	415	12	AQ228137	AQ228137 HS_2014_B
c	43	59.6	19.8	976	10	BM423178	BM423178 PLATE3_C0
	44	59.4	19.7	415	12	AQ210277	AQ210277 HS_3249_B
c	45	59.4	19.7	433	12	AQ475579	AQ475579 CITBI-E1-

ALIGNMENTS

RESULT 1
AQ703007
LOCUS AQ703007 531 bp DNA linear GSS 07-JUL-1999
DEFINITION HS_5443_B1_G09_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1019 Col=17 Row=N, DNA sequence.
ACCESSION AQ703007
VERSION AQ703007.1 GI:5412433
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 1019 row: N column: 17
Seq primer: T7
Class: BAC ends
High quality sequence stop: 531.
FEATURES Location/Qualifiers
source 1. .531
/organism="Homo sapiens"

```
                     /db_xref="taxon:9606"
                     /clone="Plate=1019 Col=17 Row=N"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      129 a    114 c    146 g    137 t      5 others
ORIGIN


  Query Match             86.4%;  Score 260.2;  DB 12;  Length 531;
  Best Local Similarity   98.1%;  Pred. No. 2.7e-55;
  Matches  262;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     264 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 323

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     324 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 383

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 443

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||| |||||||| |||||||||||| |||||||||||||||||||||||||||| |
Db     444 GCCCTNGCTTTTTATATAACCAATAATCTTNTGTATATTTATACCTGTTAAAAATTCAAA 503

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCA 267
          |||||||| ||||||||||||||||||
Db     504 AATGTCAACGCCGGGCACGGTGGCTCA 530


RESULT   2
AG019823
LOCUS       AG019823                630 bp    DNA     linear   GSS 16-OCT-1999
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: B2289H10 SpN077(-21),
            genomic survey sequence.
ACCESSION   AG019823
VERSION     AG019823.1  GI:6045767
KEYWORDS    GSS.
SOURCE      Homo sapiens DNA, clone:B2289H10 SpN077(-21).
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
            Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in DataBase (1999) In press
REFERENCE   2  (bases 1 to 630)
  AUTHORS   Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
            Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-1999) to the DDBJ/EMBL/GenBank databases.
            Masahira Hattori, RIKEN Genomic Sciences Center(GSC) c/o Kitasato
            University; 1-15-1 Kitasato, Sagamihara 228-8555, Japan
            (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:042-778-9923,
            Fax:042-778-9924)
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="B2289H10 SpN077(-21)"
                     /map="21q"
BASE COUNT      193 a    126 c    137 g    174 t
ORIGIN


  Query Match             21.7%;  Score 65.2;  DB 12;  Length 630;
  Best Local Similarity   69.8%;  Pred. No. 2.6e-06;
  Matches   88;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy     176 CTATGGCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAAT 235
           | | |  |  |  ||  ||| | ||||  ||  |||    |  ||  | |  |||||||
Db      81 CAAAGAGCTATATTTACTATATGACCAGAAAAATTTAAAGTCATTGCAACCTTTAAAAAA 140

Qy     236 TCAGAAATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGC 295
           | | |  | | | ||| ||||||||||||||||| ||||||||||||||||||||||||
Db     141 TTATAGTTATGATGGCTGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGC 200

Qy     296 CGAGGC 301
           ||||||
Db     201 TGAGGC 206


RESULT   3
AG020621
LOCUS       AG020621                659 bp    DNA     linear   GSS 19-OCT-1999
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: B266G18_SpN017_Rv,
            genomic survey sequence.
ACCESSION   AG020621
VERSION     AG020621.1  GI:6063581
KEYWORDS    GSS.
SOURCE      Homo sapiens DNA, clone:B266G18_SpN017_Rv.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
            Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in DataBase (1999) In press
REFERENCE   2  (bases 1 to 659)
  AUTHORS   Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
            Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-OCT-1999) to the DDBJ/EMBL/GenBank databases.
            Masahira Hattori, RIKEN Genomic Sciences Center(GSC) c/o Kitasato
            University; 1-15-1 Kitasato, Sagamihara 228-8555, Japan
            (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:042-778-9923,
            Fax:042-778-9924)
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="B266G18_SpN017_Rv"
                     /map="21q"
BASE COUNT      205 a    132 c    138 g    184 t
ORIGIN


  Query Match             21.7%;  Score 65.2;  DB 12;  Length 659;
  Best Local Similarity   69.8%;  Pred. No. 2.6e-06;
  Matches   88;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy     176 CTATGGCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAAT 235
           | | |  |  |  ||  ||| | ||||  ||  |||    |  ||  | |  |||||||
Db      81 CAAAGAGCTATATTTACTATATGACCAGAAAAATTTAAAGTCATTGCAACCTTTAAAAAA 140

Qy     236 TCAGAAATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGC 295
           | | |  | | | ||| ||||||||||||||||| ||||||||||||||||||||||||
Db     141 TTATAGTTATGATGGCTGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGC 200

Qy     296 CGAGGC 301
           ||||||
Db     201 TGAGGC 206
```

```
RESULT  4
W44873
LOCUS       W44873                  422 bp    mRNA    linear   EST 10-OCT-1996
DEFINITION  zc05d01.r1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:321409 5' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   W44873
VERSION     W44873.1  GI:1328963
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 597   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 350.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /db_xref="GDB:1259067"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:321409"
                     /clone_lib="Soares_parathyroid_tumor_NbHPA"
                     /tissue_type="parathyroid tumor"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTT
                     TTTTT-3'], double-stranded cDNA was size selected, ligated
                     to Eco RI adapters (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of a modified pT7T3
                     vector (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                     adenomas was kindly provided by Dr. Stephen Marx, National
                     Institute of Diabetes and Digestive and Kidney Diseases,
                     NIH."
BASE COUNT      117 a     96 c     93 g    106 t     10 others
ORIGIN
```

```
  Query Match             21.6%;  Score 65;  DB 10;  Length 422;
  Best Local Similarity   79.4%;  Pred. No. 2.9e-06;
  Matches   77;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     205 AATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGGCACGGTGGC 264
           |||  |||||   | ||  ||  ||| ||||| | |||||  | ||| || || ||||||
Db     231 AATAGTTTGTCCTTCTAGTCCCCTTATAAATTAATAAATGGAAGGGCTGGTCATGGTGGC 290

Qy     265 TCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
           ||||  ||||||||||||||||||||||||||||||||
Db     291 TCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 327
```

```
RESULT  5
BG388094/c
LOCUS       BG388094                651 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  602413054F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4521732 5',
            mRNA sequence.
ACCESSION   BG388094
VERSION     BG388094.1  GI:13281540
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 651)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10421  row: k  column: 13
            High quality sequence stop: 633.
FEATURES             Location/Qualifiers
     source          1. .651
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4521732"
                     /clone_lib="NIH_MGC_92"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      180 a    160 c    137 g    174 t
ORIGIN
```

```
  Query Match             21.3%;  Score 64;  DB 10;  Length 651;
  Best Local Similarity   70.8%;  Pred. No. 5.2e-06;
  Matches   85;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           |||  ||   | | |||  ||  |  ||  |  || || |||  ||||  || || | ||  | |
Db     366 GCCATTCTTGTATATGTATACCATCTTGATTGGTTTATATATAGATGGTATATATCTAAA 307

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           |  |  |||| || | | |||||||||  ||||||||||||||||||||||||||||  |||
Db     306 TACATATAGGCTGGACGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG 247
```

```
RESULT  6
AQ471624
LOCUS       AQ471624                533 bp    DNA     linear   GSS 23-APR-1999
DEFINITION  CITBI-E1-2590G21.TR CITBI-E1 Homo sapiens genomic clone 2590G21,
            DNA sequence.
ACCESSION   AQ471624
VERSION     AQ471624.1  GI:4655278
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 533)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
```

```
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CITBI-E1-2590G21.TF
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .533
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2590G21"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
BASE COUNT      166 a    109 c    113 g    145 t
ORIGIN


  Query Match             21.1%;  Score 63.6;  DB 12;  Length 533;
  Best Local Similarity   71.2%;  Pred. No. 6.6e-06;
  Matches   84;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy    184 CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAAT 243
          ||   ||  || | |||  | || | ||  |  |||| ||  | ||||| ||  ||  ||
Db     80 CTGCTTTAATACTAAACTTAAAAGCATTGTTTGATTTTTAAATTTTAAAGATAAAGGGAT 139

Qy    244 GTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
             |||||||||||| ||||||||  |||||||||||| |||||| ||||||||||||
Db    140 AATTAGGCCGGGCACAGTGGCTCATGCCTGTAATCCCAACACTTTAGGAGGCCGAGGC 197


RESULT   7
AZ756856/c
LOCUS       AZ756856               303 bp    DNA     linear   GSS 01-MAR-2001
DEFINITION  ew02e08.x1 PAX3/FKHR CASTing Library 'ew' Homo sapiens genomic
            clone ew02e08 random, DNA sequence.
ACCESSION   AZ756856
VERSION     AZ756856.1  GI:13176313
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 303)
  AUTHORS   Barber,T.D., Barber,M.C., Tomescu,O., Barr,F., Ruben,S. and
            Friedman,T.B.
  TITLE     Cyclic amplification and selection of target genes regulated by
            Pax3 and PAX3/FKHR in embryogenesis and alveolar rhabdomyosarcoma
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Friedman TB
            Laboratory of Molecular Genetics
            National Institute on Deafness and Other Communication Disorders,
            National Institutes of Health
            5 Research Court, Room 2A-15, Rockville, MD 20850, USA
            Tel: 301 402 7580
            Fax: 301 496 7882
            Email: friedman@nidcd.nih.gov
            Plate: 02  row: e  column: 08
            Seq primer: -21M13 forward primer (ABI)
            Class: random plasmid subclone.
FEATURES             Location/Qualifiers
     source          1. .303
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ew02e08"
                     /clone_lib="PAX3/FKHR CASTing Library 'ew'"
                     /sex="Male"
                     /lab_host="DH10B"
                     /note="Vector: pGEM-T Easy; Human genomic DNA was
                     partially digested with Sau3AI, ligated to ds  linkers,
                     and enriched for binding to human PAX3/FKHR protein using
                     a  Whole Genome PCR-based strategy.  DNA fragments
                     containing putative  PAX3/FKHR binding sites were
                     amplified by PCR and cloned into pGEM-T  Easy (Promega).
                     The ligation products were transformed into DH10B
                     electrocompetent cells (Life Technologies)."
BASE COUNT       97 a     61 c     58 g     87 t
ORIGIN


  Query Match             21.1%;  Score 63.4;  DB 12;  Length 303;
  Best Local Similarity   72.6%;  Pred. No. 7.3e-06;
  Matches   82;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy    189 TTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAA 248
          |||    ||  ||||  |   |||  ||  ||| |||  || ||    |  ||| ||
Db    126 TTTCCCATTTTCCAAGCAAACTTTACATTCTTAAACCATTTTAATTAGCGTAAAAGTATC 67

Qy    249 GGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
          |||||||| |||||||||||| ||||||||||||||||||||||||||||||
Db     66 GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 14


RESULT   8
AA825901/c
LOCUS       AA825901               389 bp    mRNA    linear   EST 09-JUN-1998
DEFINITION  od52h11.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1371621 3'
            similar to contains Alu repetitive element;contains element MER9
            repetitive element ;, mRNA sequence.
ACCESSION   AA825901
VERSION     AA825901.1  GI:2899213
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 726   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 388.
FEATURES             Location/Qualifiers
```

```
     source          1. .389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1371621"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       94 a     79 c     89 g    127 t
ORIGIN


  Query Match             20.9%;  Score 63;  DB 9;  Length 389;
  Best Local Similarity   61.1%;  Pred. No. 9.3e-06;
  Matches  102;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy     135 TTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCCCTTGCTTTTTA 194
           |||  |   |||| | ||  |  |  | | |||    |   |  |     |     | | |
Db     274 TTAGTGAGTTTTATTCAGACCAAGTGGTGAAATCTGAAAATTGAGATTTCACATGTATGA 215

Qy     195 TTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGG 254
           ||  ||   |  | |      |  |||||| ||| | ||||     |||  || | |||||
Db     214 TTCAAATGACTACCCAAATGAGTTTTATAACTGATCAAAAAATTTAAAAATCCAAGCCGG 155

Qy     255 GCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
           ||| ||||||||||| |||||||||||| ||||||||||||||||||
Db     154 GCATGGTGGCTCACACCTGTAATCCCAACACTTTGGGAGGCCGAGGC 108


RESULT  9
AI092525/c
LOCUS       AI092525                342 bp    mRNA    linear   EST 18-AUG-1998
DEFINITION  qa38f07.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1689061
            3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AI092525
VERSION     AI092525.1  GI:3431501
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 342)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .342
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1689061"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      109 a     72 c     75 g     86 t
ORIGIN


  Query Match             20.9%;  Score 62.8;  DB 9;  Length 342;
  Best Local Similarity   71.9%;  Pred. No. 1e-05;
  Matches   82;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy     188 CTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCA 247
           || |||| ||  |  ||  |   |||    |  || | |   || ||   ||| |||  |
Db     182 CTGTTTAGTTTTCAGATCTTTCCTTGCCATTCCATCCTTCCGAATAAAAGAGATATGAAA 123

Qy     248 AGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
           ||||||||||  |||||||| ||||||||||||||||||||||||||||||||
Db     122 AGGCCGGGCACAATGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 69


RESULT  10
AQ150673/c
LOCUS       AQ150673                439 bp    DNA     linear   GSS 08-OCT-1998
DEFINITION  HS_3203_A2_A05_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3203 Col=10 Row=A, DNA sequence.
ACCESSION   AQ150673
VERSION     AQ150673.1  GI:3544131
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 439)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3203  row: A  column: 10
            Class: BAC ends
            High quality sequence stop: 439.
FEATURES             Location/Qualifiers
     source          1. .439
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3203 Col=10 Row=A"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      126 a    103 c     95 g    114 t      1 others
ORIGIN


  Query Match             20.9%;  Score 62.8;  DB 12;  Length 439;
  Best Local Similarity   59.2%;  Pred. No. 1e-05;
```

```
  Matches  106;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy    122 ACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGG 181
          |||  | | |  | | | | | ||  | |  ||||        ||| ||||   || |
Db    433 ACACTTTAAGTATCTCACTGCATTCTTGCTGGCATGATTTCTGAAGAGAAGATGCTGTAA 374

Qy    182 CCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAA 241
          |||   |  ||| ||   ||   ||||  |    | ||   | |||| |||| |
Db    373 TACTTATCTCNTATGTAGGTAAGGTGCTTTCTTCCTCTGATTATTTTTAAGAATTTACTT 314

Qy    242 ATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          |  |  | | ||||| |||||||||  |||||||||||||||||||||||||| ||||
Db    313 TTACCTGGCCTGGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGG 255


RESULT  11
AW020486
LOCUS       AW020486                553 bp    mRNA    linear   EST 28-AUG-2001
DEFINITION  df10e08.y1 Morton Fetal Cochlea Homo sapiens cDNA clone
            IMAGE:2483031 5', mRNA sequence.
ACCESSION   AW020486
VERSION     AW020486.2  GI:15330169
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
            and Morton,C.C.
  TITLE     Isolation of novel and known genes from a human fetal cochlear cDNA
            library using subtractive hybridization and differential screening
  JOURNAL   Genomics 23, 42-50 (1994)
  MEDLINE   95130111
COMMENT     On Sep 13, 1999 this sequence version replaced gi:5874016.
            Contact: Morton, C. C.
            Departments of Pathology and Obstetrics, Gynecology and
            Reproductive Biology
            Brigham and Women's Hospital
            75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
            Tel: 617 732 7980
            Fax: 617 738 6996
            Email: ccmorton@bics.bwh.harvard.edu
            DNA sequencing and analyses were performed by National Institutes
            of Health Intramural Sequencing Center (NISC; see
            http://www.nisc.nih.gov).
             This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Plate: LLAM6178  row: I  column: 16
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2483031"
                     /clone_lib="Morton Fetal Cochlea"
                     /tissue_type="cochlea"
                     /dev_stage="16-22 week fetus"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
                     Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
                     unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
                     37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
                     XR Vector. Library constructed by N. Robertson, C. Morton.
                     -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT      156 a    101 c    112 g    184 t
ORIGIN

  Query Match             20.9%;  Score 62.8;  DB 9;  Length 553;
  Best Local Similarity   65.9%;  Pred. No. 1e-05;
  Matches   91;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy    164 AAAGGGAAGTGTCTATGGCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATA 223
          | ||   ||  ||||  |   ||   ||| ||||| || || ||| | || |     | |
Db    303 AGAGCTGAGGATCTAAAGATTTTTACTTTGATTTATCCCATTATCATCTGCAGGGAAACA 362

Qy    224 CCTGTTAAAAATTCAGAAATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGC 283
          || |   | |    |||  |  |||| |||||| | ||||||| ||||||||||||
Db    363 ATTGCTTTTACTGATTAAAAATGCAGGCTGGGCACAGCGGCTCACGCCTGTAATCCCAGT 422

Qy    284 ACTTTGGGAGGCCGAGGC 301
          ||||||||||||||||||
Db    423 ACTTTGGGAGGCCGAGGC 440


RESULT  12
AI565873
LOCUS       AI565873                281 bp    mRNA    linear   EST 13-MAY-1999
DEFINITION  tq49g04.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2212182 3'
            similar to contains Alu repetitive element;contains L1.t1 L1
            repetitive element ;, mRNA sequence.
ACCESSION   AI565873
VERSION     AI565873.1  GI:4524325
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 281)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 966   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 280
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .281
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2212182"
                     /clone_lib="NCI_CGAP_Ut1"
                     /tissue_type="well-differentiated endometrial
                     adenocarcinoma, 7 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Life Technologies catalog #:
                     11538-014"
BASE COUNT      101 a     53 c     58 g     69 t
ORIGIN


  Query Match             20.7%;  Score 62.2;  DB 9;  Length 281;
  Best Local Similarity   76.8%;  Pred. No. 1.5e-05;
  Matches   76;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy    203 ATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGGCACGGTG 262
          | |  | ||   |||||  |     |||||||   |||||  ||| ||||||||| |||||
Db    104 ACATACATTGAGATATTATTCAGCCTTAAAAAAGAAGAAAAATCATGGCCGGGCGCGGTG 163

Qy    263 GCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
          |||||| |||||||||||||||||||||||||||||||| |
```

```
Db    164 GCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGAC 202


RESULT  13
BF898122
LOCUS       BF898122                310 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  RC1-MT0160-281100-021-g01 MT0160 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF898122
VERSION     BF898122.1  GI:12289581
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC1&t2=RC1-MT0160-
            281100-021-g01&t3=2000-11-28&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 310.
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MT0160"
                     /dev_stage="Adult"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       86 a     64 c     59 g    101 t
ORIGIN


  Query Match           20.7%;  Score 62.2;  DB 10;  Length 310;
  Best Local Similarity   67.2%;  Pred. No. 1.5e-05;
  Matches   88;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy    170 AAGTGTCTATGGCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTT 229
          || | |  |  |  | | |  |||| || | ||||      ||   | | ||   ||| |
Db    173 AAATTTGAAATGTACATACCCTTTACTTCAGCAATTCCACTTCTAAGAATTTAGCCTATA 232

Qy    230 AAAAATTCAGAAATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTG 289
           |||  |    |  |||||||| ||| |||||||||| ||||||||||||||||||||
Db    233 GAAATATTCACAGATTCAAGGCCAGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTG 292

Qy    290 GGAGGCCGAGG 300
          ||||||||||
Db    293 GGAGGCCGAGG 303


RESULT  14
R89296
LOCUS       R89296                  400 bp    mRNA    linear   EST 24-AUG-1995
DEFINITION  yq02f11.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:195789 3' similar to contains Alu repetitive element;contains
            L1 repetitive element ;, mRNA sequence.
ACCESSION   R89296
VERSION     R89296.1  GI:954123
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 400)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1115
            High quality sequence stops: 268
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1115   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 268.
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /db_xref="GDB:3764839"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:195789"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      119 a     83 c    112 g     81 t      5 others
ORIGIN


  Query Match           20.7%;  Score 62.2;  DB 10;  Length 400;
  Best Local Similarity   76.8%;  Pred. No. 1.5e-05;
  Matches   76;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy    203 ATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGGCACGGTG 262
          | |  | ||   |||||  |      |||||||   |||||  ||| ||||||||| |||||
Db     82 ACATACATTGAGATATTATTCAGCCTTAAAAAAGAAGAAAAATCATGGCCGGGCGCGGTG 141

Qy    263 GCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
          |||||| |||||||||||||||||||||||||||||| |
Db    142 GCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGAC 180


RESULT  15
H71349
```

```
LOCUS       H71349                  408 bp    mRNA    linear   EST 26-OCT-1995
DEFINITION  yu55h06.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:230075 3' similar to contains Alu repetitive element;contains
            L1 repetitive element ;, mRNA sequence.
ACCESSION   H71349
VERSION     H71349.1  GI:1043165
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1145
            High quality sequence stops: 213
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1145   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 213.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Homo sapiens"
                     /db_xref="GDB:3781171"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:230075"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      121 a     75 c    103 g     96 t     13 others
ORIGIN


  Query Match           20.7%;  Score 62.2;  DB 10;  Length 408;
  Best Local Similarity  76.8%;  Pred. No. 1.5e-05;
  Matches   76;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     203 ATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGGCACGGTG 262
           | |  | ||   |||||  |     |||||||   |||||  ||| |||||||| |||||
Db     121 ACATACATTGAGATATTATTCAGCCTTAAAAAAGAAGAAAAATCATGGCCGGGCGCGGTG 180

Qy     263 GCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
           |||||| ||||||||||||||||||||||||||||||| |
Db     181 GCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGAC 219


Search completed: June 19, 2002, 10:24:39
```

Job time: 7196 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:54:02 ; Search time 213.18 Seconds
(without alignments)
2424.201 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870
Perfect score: 301
Sequence: 1 GGCACGGAATCCCTGGTTGG.....GCACTTTGGGAGGCCGAGGC 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	301	100.0	10825	18	AAT96690	Hereditary haemoch
	2	301	100.0	10825	22	AAC68425	Human hereditary h
	3	301	100.0	10825	22	AAC68426	Human hereditary h
	4	301	100.0	10825	22	AAC68427	Human hereditary h
	5	301	100.0	10825	22	AAC68428	Human hereditary h
	6	301	100.0	12146	21	AAA96794	Genomic DNA of a h
	7	299.4	99.5	5749	22	AAL36747	Human musculoskele
c	8	294.6	97.9	237326	19	AAV57903	Hereditary haemoch
c	9	61.4	20.4	429	21	AAH30227	Human colon cancer
	10	61.4	20.4	26110	22	AAK65036	Human immune/haema
	11	61.4	20.4	26110	22	AAK78526	Human immune/haema
	12	61	20.3	584	22	AAK79701	Human immune/haema
	13	61	20.3	6534	22	AAL02766	Human reproductive
	14	60.6	20.1	344	21	AAC27863	Human secreted pro
c	15	60.6	20.1	3242	22	AAI86495	Human polynucleoti
c	16	60.6	20.1	5768	22	AAL05331	Human reproductive
	17	60.6	20.1	7619	22	AAS44934	cDNA encoding nove
c	18	60	19.9	655	22	AAK56970	Human immune/haema
	19	60	19.9	42519	22	AAK81318	Human immune/haema
c	20	59.8	19.9	448	22	AAI93368	Human polynucleoti
c	21	59.6	19.8	841	22	AAK85928	Human immune/haema
	22	59.6	19.8	22655	22	AAK70122	Human immune/haema
	23	59.2	19.7	16595	20	AAX23521	Human kidney amino
c	24	59	19.6	606	22	AAL04682	Human reproductive
c	25	59	19.6	5270	22	AAK64901	Human immune/haema
	26	58.8	19.5	328	21	AAC30053	Human secreted pro
c	27	58.8	19.5	4156	22	AAK83412	Human immune/haema
	28	58.6	19.5	1146	20	AAX85051	Human secreted pro
	29	58.6	19.5	4512	22	AAH16662	Human cDNA sequenc
	30	58.6	19.5	17979	22	AAK64964	Human immune/haema
	31	58.6	19.5	17979	22	AAK71664	Human immune/haema
	32	58.4	19.4	270	22	AAK56284	Human immune/haema
	33	58.4	19.4	338	22	AAS38872	Novel human diagno
	34	58.4	19.4	484	22	AAI93692	Human polynucleoti
c	35	58.4	19.4	697	22	AAH70948	Human cervical can
c	36	58.4	19.4	986	21	AAC57483	Arachidonic acid m
	37	58.4	19.4	1641	22	ABA16008	Human nervous syst
c	38	58.4	19.4	6004	22	AAK84542	Human immune/haema
	39	58.4	19.4	14969	22	AAK78763	Human immune/haema
c	40	58.4	19.4	21837	22	AAK85946	Human immune/haema
	41	58.4	19.4	57728	22	AAC87588	Human 9p11 chromos
	42	58.2	19.3	710	22	AAH98518	Human EST-derived
c	43	58.2	19.3	32184	22	AAL05850	Human reproductive
c	44	58.2	19.3	32204	22	AAL05849	Human reproductive
	45	58.2	19.3	51474	22	AAF97846	Human neuroblastom

ALIGNMENTS

RESULT 1
AAT96690
ID AAT96690 standard; DNA; 10825 BP.
XX
AC AAT96690;
XX
DT 14-APR-1998 (first entry)
XX
DE Hereditary haemochromatosis gene.
XX
KW Hereditary haemochromatosis; metal toxicity; diagnosis;
KW gene therapy; prenatal screening; human; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	361..7147
FT		/*tag= a
FT		/note= "contains introns"
FT	intron	437..3761
FT		/*tag= b
FT		/number= 1
FT	intron	4026..4234
FT		/*tag= c
FT		/number= 2
FT	intron	4511..5605
FT		/*tag= d
FT		/number= 3
FT	intron	5882..6039
FT		/*tag= e
FT		/number= 4
FT	intron	6154..7106

```
FT                  /*tag=  f
FT                  /number= 5
FT   mutation       3872
FT                  /*tag=  g
FT                  /note= "C to G substitution (24d2 mutation)
FT                         results in His to Asp substitution"
FT   variation      3878
FT                  /*tag=  h
FT                  /note= "A to T substitution (24d7 variant)
FT                         results in Ser to Cys substitution"
FT   mutation       5834
FT                  /*tag=  i
FT                  /note= "G to A substitution (24d1 mutation
FT                         associated with HH), results in Cys to
FT                         Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   P-PSDB; AAW36499.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 3; 115pp; English.
XX
CC   This genomic DNA sequence corresponds to the human gene whose
CC   mutated form is associated with hereditary haemochromatosis (HH).
CC   To identify this novel gene, allelic association patterns were
CC   determined between known markers and the HH locus in the HLA region
CC   of chromosome 6.  A physical clone coverage was then generated
CC   extending from D6S265, which is a marker that is centromeric of
CC   HLA-A, in a telomeric direction through D6S276, a marker at which
CC   the allelic association was no longer observed.  A single mutation
CC   (24d1) in the HH gene appears responsible for the majority of HH
CC   disease.  This comprises a G to A substitution that is present in
CC   86% of affected chromosomes and in 4% of unaffected chromosomes.
CC   It results in a Cys to Tyr substitution in the encoded protein (see
CC   AAW36499) at a critical disulphide bridge important for secondary
CC   structure.  The following are claimed: the HH genomic DNA (I), a
CC   1437 bp cDNA sequence (Ia) (see AAT96691) and their 24d1, 24d2 and
CC   24d7 variants; a cloning or expression vector; host cells; a
CC   peptide product chosen from the HH gene product, its variants
CC   (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC   residues of these; an antibody produced using the peptide; a method
CC   to determine the presence or absence of the common HH gene
CC   mutation; an animal model for the HH disease; metal chelation
CC   agents, T-cell differentiation factors and therapeutic agents for
CC   the mitigation of injury due to oxidative process in vivo or
CC   mitigation of iron overload; a method for screening potential
CC   therapeutic agents for activity in connection with HH disease; an
CC   antisense oligonucleotide directed against a transcriptional
CC   product of a nucleic acid sequence as above; and oligonucleotides
CC   or pairs of oligonucleotides covering a range of nucleotides from
CC   (I), (Ia) or their variants, useful for detecting a polymorphism in
CC   the HH gene.  The invention also relates to methods for screening
CC   for HH homozygotes, to HH diagnosis, prenatal screening and
CC   diagnosis, and therapies of HH disease, including gene therapy,
CC   protein- and antibody-based therapeutics, and small molecule
CC   therapeutics.
XX
SQ   Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;


  Query Match             100.0%;  Score 301;  DB 18;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1e-79;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 gcccttgctttttatttaaccaataatcttttgtatatttatacctgttaaaaattcaga 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 aatgtcaaggccgggcacggtggctcacccctgtaatcccagcactttgggaggccgagg 4869

Qy     301 C 301
          |
Db    4870 c 4870


RESULT   2
AAC68425
ID   AAC68425 standard; DNA; 10825 BP.
XX
AC   AAC68425;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36869.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
```

CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;

Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 gcccttgcttttttatttaaccaataatctttttgtatatttatacctgttaaaaattcaga 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 aatgtcaaggccgggcacggtggctcaccccctgtaatcccagcactttgggaggccgagg 4869

Qy     301 C 301
          |
Db    4870 c 4870
```

RESULT 3
AAC68426
ID AAC68426 standard; DNA; 10825 BP.
XX
AC AAC68426;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36870.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;

Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 gcccttgcttttttatttaaccaataatctttttgtatatttatacctgttaaaaattcaga 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 aatgtcaaggccgggcacggtggctcaccccctgtaatcccagcactttgggaggccgagg 4869

Qy     301 C 301
          |
Db    4870 c 4870
```

RESULT 4
AAC68427
ID AAC68427 standard; DNA; 10825 BP.
XX
AC AAC68427;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d2 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX

DR WPI; 2001-006341/01.
DR P-PSDB; AAB36871.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;

Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 gcccttgcttttttatttaaccaataatcttttgtatatttatacctgttaaaaattcaga 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 aatgtcaaggccgggcacggtggctcaccccTgtaatcccagcactttgggaggccgagg 4869

Qy     301 C 301
          |
Db    4870 c 4870
```

RESULT 5
AAC68428
ID AAC68428 standard; DNA; 10825 BP.
XX
AC AAC68428;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 26-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX

PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36872.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2252 C; 2648 G; 2926 T; 0 other;

Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 gcccttgcttttttatttaaccaataatcttttgtatatttatacctgttaaaaattcaga 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 aatgtcaaggccgggcacggtggctcaccccTgtaatcccagcactttgggaggccgagg 4869

Qy     301 C 301
          |
Db    4870 c 4870
```

RESULT 6
AAA96794
ID AAA96794 standard; cDNA; 12146 BP.
XX
AC AAA96794;
XX
DT 19-FEB-2001 (first entry)
XX
DE Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX
KW Human; histocompatibility iron loading protein; HFE protein;
KW major histocompatibility complex; non-classical class I gene;
KW chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT exon 1028..1324
FT /*tag= a
FT /number= 1

```
FT   intron          1325..4651
FT                   /*tag= b
FT                   /number= 1
FT   exon            4652..4915
FT                   /*tag= c
FT                   /number= 2
FT   intron          4916..5124
FT                   /*tag= d
FT                   /number= 2
FT   exon            5125..5400
FT                   /*tag= e
FT                   /number= 3
FT   intron          5401..6493
FT                   /*tag= f
FT                   /number= 3
FT   exon            6494..6769
FT                   /*tag= g
FT                   /number= 4
FT   intron          6770..6927
FT                   /*tag= h
FT                   /number= 4
FT   exon            6928..7041
FT                   /*tag= i
FT                   /number= 5
FT   intron          7042..7994
FT                   /*tag= j
FT                   /number= 5
FT   exon            7995..9050
FT                   /*tag= k
FT                   /number= 6
FT   intron          9051..10205
FT                   /*tag= l
FT                   /number= 6
FT   exon            10206..10637
FT                   /*tag= m
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Example 1; Page 21-28; 55pp; English.
XX
CC   The present sequence represents the human histocompatibility iron
CC   loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;
```

```
  Query Match             100.0%;  Score 301;  DB 21;  Length 12146;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-79;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5460 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 5519

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5520 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 5579

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5580 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 5639

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5640 gcccttgctttttatttaaccaataatctttttgtatatttatacctgttaaaaattcaga 5699

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5700 aatgtcaaggccgggcacggtggctcacccctgtaatcccagcactttgggaggccgagg 5759

Qy     301 C 301
          |
Db    5760 c 5760
```

```
RESULT   7
AAL36747
ID   AAL36747 standard; DNA; 5749 BP.
XX
AC   AAL36747;
XX
DT   08-JAN-2002  (first entry)
XX
DE   Human musculoskeletal system related polynucleotide SEQ ID NO 3112.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein;
KW   musculoskeletal system; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155367-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01338.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
```

PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451937/48.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
PS Example 2; SEQ ID NO 3112; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and

```
CC   parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5749 BP; 1600 A; 1192 C; 1403 G; 1553 T; 1 other;


  Query Match             99.5%;  Score 299.4;  DB 22;  Length 5749;
  Best Local Similarity   99.7%;  Pred. No. 2.5e-79;
  Matches  300;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     570 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     630 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     690 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     750 gcccttgcttttttatttaaccaataatctttttgtatatttatacctgttaaaaattcaga 809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     810 aatgtcaaggccgggcacggtggctcacccttgtaatcccagcactttgggaggccgagg 869

Qy     301 C 301
          |
Db     870 c 870


RESULT   8
AAV57903/c
ID   AAV57903 standard; DNA; 237326 BP.
XX
AC   AAV57903;
XX
DT   21-DEC-1998  (first entry)
XX
DE   Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW   Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW   diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW   BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW   type 1 sodium transport gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO9814466-A1.
XX
PD   09-APR-1998.
XX
PF   30-SEP-1997;   97WO-US17658.
XX
PR   07-MAY-1997;   97US-0852495.
PR   01-OCT-1996;   96US-0724394.
XX
PA   (PROG-) PROGENTIOR INC.
XX
PI   Feder JN,  Kronmal GS,  Lauer PM,  Ruddy DA,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1998-240014/21.
XX
PT   Hereditary haemochromatosis gene products - used to develop products
PT   for the diagnosis and treatment of hereditary disorders in iron
PT   metabolism
XX
PS   Claim 1; Fig 9; 209pp; English.
XX
CC   The present invention describes hereditary haemochromatosis gene
CC   products from the human haemochromatosis gene. The present sequence
CC   represents a hereditary haemochromatosis subregion from an hereditary
CC   haemochromatosis (HH) affected individual. Also described is a
CC   method to determine the presence or absence of the common hereditary
CC   haemochromatosis (HFE) gene mutation in an individual comprising:
CC   (a) providing DNA or RNA from the individual; and (b) assessing the
CC   DNA or RNA for the presence or absence of a haplotype or genotype where
CC   the presence or absence of the haplotype genotype indicates the likely
CC   presence of the HFE gene mutation in the genome of the individual. The
CC   HFE gene sequences from the present invention can be used to develop
CC   products for use in the diagnosis and treatment of HFE. The present
CC   invention also describes BTF genes, which are homologues of the milk
CC   protein butyrophilin (BT), and can be used in the production of agonists
CC   and antagonists of BT function. Also described are: (1) a RoRet gene
CC   which can be used to develop products for the study, diagnosis and
CC   treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC   which are homologues of a type 1 sodium transport gene, and can
CC   similarly be used for hypophosphatemia.
XX
SQ   Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;


  Query Match             97.9%;  Score 294.6;  DB 19;  Length 237326;
  Best Local Similarity   98.7%;  Pred. No. 2.7e-77;
  Matches  297;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42530 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 42471

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42470 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 42411

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42410 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 42351

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42350 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 42291

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          |||||||||||||||||||  |  ||||||||||||||||||||||||||||||||||||
Db   42290 AATGTCAAGGCCGGGCACGGCCGGGCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 42231

Qy     301 C 301
          |
Db   42230 C 42230


RESULT   9
AAH30227/c
ID   AAH30227 standard; cDNA; 429 BP.
XX
AC   AAH30227;
XX
DT   27-JUL-2001  (first entry)
XX
DE   Human colon cancer cell line Km12L4-A cDNA library derived sequence #161.
XX
KW   Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW   detection; colon cancer cell line Km12L4-A; ss.
XX
OS   Homo sapiens.
XX
PN   WO200018916-A2.
XX
PD   06-APR-2000.
```

```
XX
PF   23-SEP-1999;   99WO-US22226.
XX
PR   28-SEP-1998;   98US-0102161.
PR   28-SEP-1998;   98US-0102180.
PR   29-SEP-1998;   98US-0102380.
PR   08-OCT-1998;   98US-0103815.
PR   27-OCT-1998;   98US-0105877.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT,  Escobedo J,  Innis MA,  Garcia PD,  Sudduth-Klinger J;
PI   Reinhard C,  Giese K,  Randazzo F,  Kennedy GC,  Pot D,  Kassam A;
PI   Lamson G,  Drmanac R,  Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I;
PI   Leshkowitz D,  Kita D,  Garcia V,  Jones LW,  Stache-Crain B;
XX
DR   WPI; 2000-293155/25.
XX
PT   Polynucleotide library comprising 1079 defined sequences,  useful in
PT   the form of an array to detect cancer or susceptibility to cancer -
XX
PS   Claim 1; Page 229-230; 502pp; English.
XX
CC   The present invention describes a library of polynucleotides comprising
CC   1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
CC   are: (1) an isolated polynucleotide (I) having at least 90% identity to
CC   one of the 1079 sequences; (2) a recombinant host cell containing (I);
CC   (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that
CC   specifically binds to (II); (5) a vector comprising (I); and (6) a method
CC   of detecting differentially expressed genes correlated with a cancerous
CC   state of a mammalian cell comprising detecting a gene product encoded by
CC   65 of the 1079 sequences given in the specification. The polynucleotides
CC   are used to monitor patients having (or susceptible) to cancer to detect
CC   potentially malignant events at a molecular level before they are
CC   detectable at a gross morphological level. The polynucleotides are also
CC   useful for monitoring the efficacy of various therapies and preventive
CC   interventions. Polynucleotide probes based on the disclosed sequences
CC   are useful for chromosome mapping and detection of transcription levels.
CC   The 1079 polynucleotide sequences were derived from a human colon cancer
CC   cell line Km12L4-A cDNA library.
XX
SQ   Sequence 429 BP; 142 A; 75 C; 75 G; 113 T; 24 other;


  Query Match             20.4%;  Score 61.4;  DB 21;  Length 429;
  Best Local Similarity   69.7%;  Pred. No. 1.1e-08;
  Matches   83;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy    183 CCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAA 242
          | ||  ||| |||||    ||   ||| |  ||   ||||   |||  ||||   | ||
Db    143 CTTTTTTTTGTATTTTCTAAAATGTCTATAATAATGTTATTTTTGTCCAAAAAGAAAAAT 84

Qy    243 TGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
            | | ||||||||  |||||||||  |||||||||  ||||||||||||||||||||
Db     83 GTTTATGGCCGGGCGTGGTGGCTCAAACCTGTAATCTCAGCACTTTGGGAGGCCGAGGC 25


RESULT  10
AAK65036
ID   AAK65036 standard; DNA; 26110 BP.
XX
AC   AAK65036;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:19848.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
```

PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;

XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 19848; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 26110 BP; 7464 A; 4942 C; 5567 G; 8137 T; 0 other;

Query Match 20.4%; Score 61.4; DB 22; Length 26110;
Best Local Similarity 72.1%; Pred. No. 5.3e-08;
Matches 80; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 190 TTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAG 249
Db 8013 tctgttttatccaacttttaatgttctttcaaatcctcttaaaactaagaatttttagg 8072

Qy 250 GCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
Db 8073 gctggacacggtggctcacgcctgtaatcccagcactctgggaggctgagg 8123

RESULT 11
AAK78526
ID AAK78526 standard; DNA; 26110 BP.
XX
AC AAK78526;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33338.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 33338; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased

```
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 26110 BP; 7464 A; 4942 C; 5567 G; 8137 T; 0 other;


  Query Match             20.4%;  Score 61.4;  DB 22;  Length 26110;
  Best Local Similarity   72.1%;  Pred. No. 5.3e-08;
  Matches   80;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy     190 TTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAG 249
           | |  |||| ||||   | |  |||   || | | |   | |||| | |||| | | | |
Db    8013 tctgttttatccaacttttaatgttctttcaaatcctcttaaaactaagaatttttagg 8072

Qy     250 GCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           || || |||||||||||| |||||||||||||||| |||||||| ||||
Db    8073 gctggacacggtggctcacgcctgtaatcccagcactctgggaggctgagg 8123


RESULT  12
AAK79701
ID   AAK79701 standard; DNA; 584 BP.
XX
AC   AAK79701;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34513.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
```

```
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 34513; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 584 BP; 196 A; 109 C; 135 G; 144 T; 0 other;

  Query Match             20.3%;  Score 61;  DB 22;  Length 584;
  Best Local Similarity   56.9%;  Pred. No. 1.7e-08;
  Matches  112;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy     105 AGATGTATGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCA 164
           ||||  |  |   ||  |  | |||||     || | |  | |  |    | ||||  |
Db     155 agataaacaaaggaggaaagattcatgtaagaaaatgccatgcaactacaaaatgggata 214

Qy     165 AAGGGAAGTGTCTATGGCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATAC 224
           ||   || ||||     | ||   ||   |  |   | |   | | || |    |   |
Db     215 aaatcaattgtcgtatacactgatttagaaaattctccaatgtatattattacatggaaa 274

Qy     225 CTGTTAAAAATTCAGAAATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCA 284
               ||||| |   |||      ||||||||||||||||||| ||||||||| ||||
Db     275 aaagcaaaaagttccaaaatattgggccgggcacggtggctcactcctgtaatcctagca 334

Qy     285 CTTTGGGAGGCCGAGGC 301
           |||||||||||||||||
Db     335 ctttgggaggccgaggc 351


RESULT  13
AAL02766
ID   AAL02766 standard; DNA; 6534 BP.
XX
AC   AAL02766;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen DNA SEQ ID NO: 5454.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
```

PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
PS Disclosure; SEQ ID NO 5454; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX
SQ Sequence 6534 BP; 1707 A; 1511 C; 1672 G; 1644 T; 0 other;

Query Match 20.3%; Score 61; DB 22; Length 6534;
Best Local Similarity 60.6%; Pred. No. 4.1e-08;
Matches 100; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 137 AATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCCCTTGCTTTTTATT 196
Db 2861 aaattatttagtggctacttgtggttgtagagtaaggtccaaattccaaagcctttaaaa 2920

Qy 197 TAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGGC 256
Db 2921 ttcacagcacattcagctacacttgagtcctcactattttaagaataggctgtgccgggc 2980

Qy 257 ACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
Db 2981 atggtggctcacgcctgtaatcccagcactttgggaggccgaggc 3025

RESULT 14
AAC27863
ID AAC27863 standard; cDNA; 344 BP.

```
XX
AC   AAC27863;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 31938.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 31938; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 344 BP; 96 A; 79 C; 83 G; 85 T; 1 other;


  Query Match             20.1%;  Score 60.6;  DB 21;  Length 344;
  Best Local Similarity   70.4%;  Pred. No. 1.8e-08;
  Matches   81;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy     186 TGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGT 245
           | | ||  | ||  |    | | || |  || |  |||  ||  ||| |   || |||
Db     137 ttccttgcagtttccacccatttttcttaatgtgaataattgggaaagaaaaaggaatca 196

Qy     246 CAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           |||||||||||||||||||  |||||||||||||||||||||||||| ||||
Db     197 taaggccgggcacggtggctcatacctgtaatcccagcactttgggaggctgagg 251


RESULT  15
AAI86495/c
ID   AAI86495 standard; cDNA; 3242 BP.
XX
AC   AAI86495;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 6555.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04927.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   P-PSDB; AA006564.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing
PT   diagnosing and treating e.g. leukaemia, inflammation and immune
PT   disorders -
XX
PS   Claim 1; SEQ ID NO 6555; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3242 BP; 949 A; 656 C; 666 G; 971 T; 0 other;


  Query Match             20.1%;  Score 60.6;  DB 22;  Length 3242;
  Best Local Similarity   79.1%;  Pred. No. 4.2e-08;
  Matches   72;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     211 TTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGGCACGGTGGCTCACCC 270
           ||||||   ||    |   ||| ||   |||  || || ||| ||||||||||||||| |
Db     328 TTGTATTGCTACTGCAACTAACAAAAAAGATGTGGCAGGGCTGGGCACGGTGGCTCACGC 269

Qy     271 CTGTAATCCCAGCACTTTGGGAGGCCGAGGC 301
           |||||||||||||||||||||||||||||||
Db     268 CTGTAATCCCAGCACTTTGGGAGGCCGAGGC 238


Search completed: June 19, 2002, 11:01:22
Job time: 7640 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:41:02 ; Search time 46.2 Seconds
(without alignments)
1600.339 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870
Perfect score: 301
Sequence: 1 GGCACGGAATCCCTGGTTGG.....GCACTTTGGGAGGCCGAGGC 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	301	100.0	10825	3	US-08-652-265-1	Sequence 1, Appli
2	301	100.0	10825	3	US-08-652-265-3	Sequence 3, Appli
3	301	100.0	10825	3	US-08-652-265-5	Sequence 5, Appli
4	301	100.0	10825	3	US-08-652-265-7	Sequence 7, Appli
5	301	100.0	10825	3	US-08-834-497A-1	Sequence 1, Appli
6	301	100.0	10825	3	US-08-834-497A-3	Sequence 3, Appli
7	301	100.0	10825	3	US-08-834-497A-5	Sequence 5, Appli
8	301	100.0	10825	3	US-08-834-497A-7	Sequence 7, Appli
9	301	100.0	10825	4	US-09-503-444A-1	Sequence 1, Appli
10	301	100.0	10825	4	US-09-503-444A-3	Sequence 3, Appli
11	301	100.0	10825	4	US-09-503-444A-5	Sequence 5, Appli
12	301	100.0	10825	4	US-09-503-444A-7	Sequence 7, Appli
13	301	100.0	12146	4	US-09-277-457-27	Sequence 27, Appl
14	301	100.0	246240	2	US-08-724-394A-20	Sequence 20, Appl
15	301	100.0	246240	2	US-08-724-394A-21	Sequence 21, Appl
16	301	100.0	246240	2	US-08-724-394A-22	Sequence 22, Appl
17	58	19.3	87350	3	US-08-781-891-79	Sequence 79, Appl
18	57.2	19.0	56516	2	US-08-996-306-1	Sequence 1, Appli
19	57.2	19.0	56516	4	US-09-338-907-1	Sequence 1, Appli
20	57.2	19.0	56516	4	US-09-218-207-1	Sequence 1, Appli
21	57.2	19.0	56520	4	US-09-338-907-179	Sequence 179, App
22	57.2	19.0	56520	4	US-09-218-207-179	Sequence 179, App
23	55.4	18.4	7210	2	US-08-257-963B-10	Sequence 10, Appl
24	55.4	18.4	7210	4	US-08-367-841A-10	Sequence 10, Appl
25	55.4	18.4	7210	5	PCT-US95-07201-10	Sequence 10, Appl
26	55.4	18.4	22481	4	US-08-367-841A-43	Sequence 43, Appl
27	55.4	18.4	22481	5	PCT-US95-07201-43	Sequence 43, Appl
28	55.2	18.3	780	4	US-09-385-982-32	Sequence 32, Appl
29	55	18.3	2808	3	US-08-870-126-7	Sequence 7, Appli
30	54.8	18.2	14796	4	US-08-975-080-35	Sequence 35, Appl
31	54.8	18.2	14796	4	US-09-630-706-10	Sequence 10, Appl
32	54.8	18.2	14796	4	US-09-496-694B-3	Sequence 3, Appli
c 33	54.2	18.0	152331	3	US-09-128-155-16	Sequence 16, Appl
c 34	54.2	18.0	176373	3	US-09-128-155-17	Sequence 17, Appl
c 35	54	17.9	3867	4	US-09-347-114A-81	Sequence 81, Appl
c 36	54	17.9	9734	4	US-09-347-114A-80	Sequence 80, Appl
c 37	53.8	17.9	8409	4	US-09-167-681-37	Sequence 37, Appl
38	53.6	17.8	1252	4	US-09-305-384-7	Sequence 7, Appli
39	53.6	17.8	1947	2	US-08-989-925-2	Sequence 2, Appli
40	53.6	17.8	2555	2	US-08-960-022-15	Sequence 15, Appl
41	53.6	17.8	6235	4	US-09-305-384-5	Sequence 5, Appli
42	53.6	17.8	6679	4	US-09-305-384-1	Sequence 1, Appli
43	53.4	17.7	2784	1	US-08-471-454-1	Sequence 1, Appli
44	53.4	17.7	2784	2	US-08-466-974-1	Sequence 1, Appli
45	53.4	17.7	2784	2	US-08-471-453-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-652-265-1
; Sequence 1, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
;   APPLICANT: Thomas, Winston J.
;   APPLICANT: Drayna, Dennis T.
;   APPLICANT: Feder, John N.
;   APPLICANT: Gnirke, Andreas
;   APPLICANT: Ruddy, David
;   APPLICANT: Tsuchihashi, Zenta
;   APPLICANT: Wolff, Roger K.
;   TITLE OF INVENTION: Hereditary Hemochromatosis Gene
;   NUMBER OF SEQUENCES: 44
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, Eighth Floor
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/652,265
;     FILING DATE: 23-MAY-1996
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Smith, William M.
;     REGISTRATION NUMBER: 30,223
;     REFERENCE/DOCKET NUMBER: 17957-000500
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 576-0200
;     TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 10825 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
;     LOCATION: 6040..6153, 7107..7147)
;     OTHER INFORMATION: /product= "Hereditary Hemochromatosis
;     OTHER INFORMATION:
```

```
;       OTHER INFORMATION:  /note= "No. 6025130mal or wild-type (unaffected)
;       OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;       OTHER INFORMATION:  allele"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) allele
;       OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;       OTHER INFORMATION:  allele (SEQ ID NO:41)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d1
US-08-652-265-1


  Query Match             100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
          |
Db    4870 C 4870
```

```
RESULT   2
US-08-652-265-3
; Sequence 3, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
```

```
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d1
US-08-652-265-3


  Query Match             100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870


RESULT   3
US-08-652-265-5
; Sequence 5, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-652-265-5


  Query Match             100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870


RESULT   4
US-08-652-265-7
; Sequence 7, Application US/08652265
```

```
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-7


  Query Match             100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870


RESULT   5
US-08-834-497A-1
; Sequence 1, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
```

```
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /note= "No. 6140305mal or wild-type (unaffected)
;      OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;      OTHER INFORMATION:  allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) allele
;      OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;      OTHER INFORMATION:  allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;      OTHER INFORMATION:  allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3878, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-1
```

```
  Query Match             100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
          |
Db    4870 C 4870
```

```
RESULT   6
US-08-834-497A-3
; Sequence 3, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-3


  Query Match           100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity  100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
           |
Db    4870 C 4870


RESULT   7
US-08-834-497A-5
; Sequence 5, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
```

```
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  genomic sequence surrounding variant
;       OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  genomic sequence surrounding variant
;       OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "g")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d2
US-08-834-497A-5


  Query Match             100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
          |
Db    4870 C 4870


RESULT   8
US-08-834-497A-7
; Sequence 7, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
```

```
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "a")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d1
US-08-834-497A-7


  Query Match             100.0%;  Score 301;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
          |
Db    4870 C 4870


RESULT   9
US-09-503-444A-1
; Sequence 1, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /note= "No. 6228594mal or wild-type (unaffected)
;      OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;      OTHER INFORMATION:  allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) allele
;      OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;      OTHER INFORMATION:  allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;      OTHER INFORMATION:  allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3878, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-1


  Query Match             100.0%;  Score 301;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870


RESULT  10
US-09-503-444A-3
; Sequence 3, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507:.6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-3


  Query Match             100.0%;  Score 301;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870


RESULT  11
US-09-503-444A-5
; Sequence 5, Application US/09503444A
```

```
; Patent No. 6228594
; GENERAL INFORMATION:
;   APPLICANT:  Thomas, Winston J.
;   APPLICANT:  Drayna, Dennis T.
;   APPLICANT:  Feder, John N.
;   APPLICANT:  Gnirke, Andreas
;   APPLICANT:  Ruddy, David
;   APPLICANT:  Tsuchihashi, Zenta
;   APPLICANT:  Wolff, Roger K.
;   TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;   NUMBER OF SEQUENCES:  44
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Pennie & Edmonds LLP
;     STREET:  1155 Avenue of the Americas
;     CITY:  New York
;     STATE:  New York
;     COUNTRY:  USA
;     ZIP:  10036
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  Windows 95
;     SOFTWARE:  WordPerfect Version 8
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/503,444A
;     FILING DATE:  14-Feb-2000
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/652,265
;     FILING DATE:  23-May-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/632,673
;     FILING DATE:  16-Apr-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/630,912
;     FILING DATE:  04-Apr-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Poissant, Brian M.
;     REGISTRATION NUMBER:  28,462
;     REFERENCE/DOCKET NUMBER:  8907-0088-999
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  212-790-9090
;     TELEFAX:  212-869-9741
;     TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  5:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  10825 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;     LOCATION:  6040..6153, 7107..7147)
;     OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;     OTHER INFORMATION:
;     OTHER INFORMATION:  mutation"
;     OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;     OTHER INFORMATION:  gene 24d2 allele"
;   FEATURE:
;     NAME/KEY:  -
;     LOCATION:  140..7319
;     OTHER INFORMATION:  /note= "start and stop positions for
;     OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;   FEATURE:
;     NAME/KEY:  -
;     LOCATION:  3852..3891
;     OTHER INFORMATION:  /note= "start and stop positions for
;     OTHER INFORMATION:  genomic sequence surrounding variant
;     OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;   FEATURE:
;     NAME/KEY:  -
;     LOCATION:  5507..6023
;     OTHER INFORMATION:  /note= "start and stop positions for
;     OTHER INFORMATION:  genomic sequence surrounding variant
;     OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;   FEATURE:
;     NAME/KEY:  allele
;     LOCATION:  replace(3872, "g")
;     OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;     OTHER INFORMATION:
;     OTHER INFORMATION:  /label= 24d2
US-09-503-444A-5


  Query Match             100.0%;  Score 301;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy    301 C 301
          |
Db   4870 C 4870


RESULT  12
US-09-503-444A-7
; Sequence 7, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
;   APPLICANT:  Thomas, Winston J.
;   APPLICANT:  Drayna, Dennis T.
;   APPLICANT:  Feder, John N.
;   APPLICANT:  Gnirke, Andreas
;   APPLICANT:  Ruddy, David
;   APPLICANT:  Tsuchihashi, Zenta
;   APPLICANT:  Wolff, Roger K.
;   TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;   NUMBER OF SEQUENCES:  44
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Pennie & Edmonds LLP
;     STREET:  1155 Avenue of the Americas
;     CITY:  New York
;     STATE:  New York
;     COUNTRY:  USA
;     ZIP:  10036
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  Windows 95
;     SOFTWARE:  WordPerfect Version 8
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/503,444A
;     FILING DATE:  14-Feb-2000
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
```

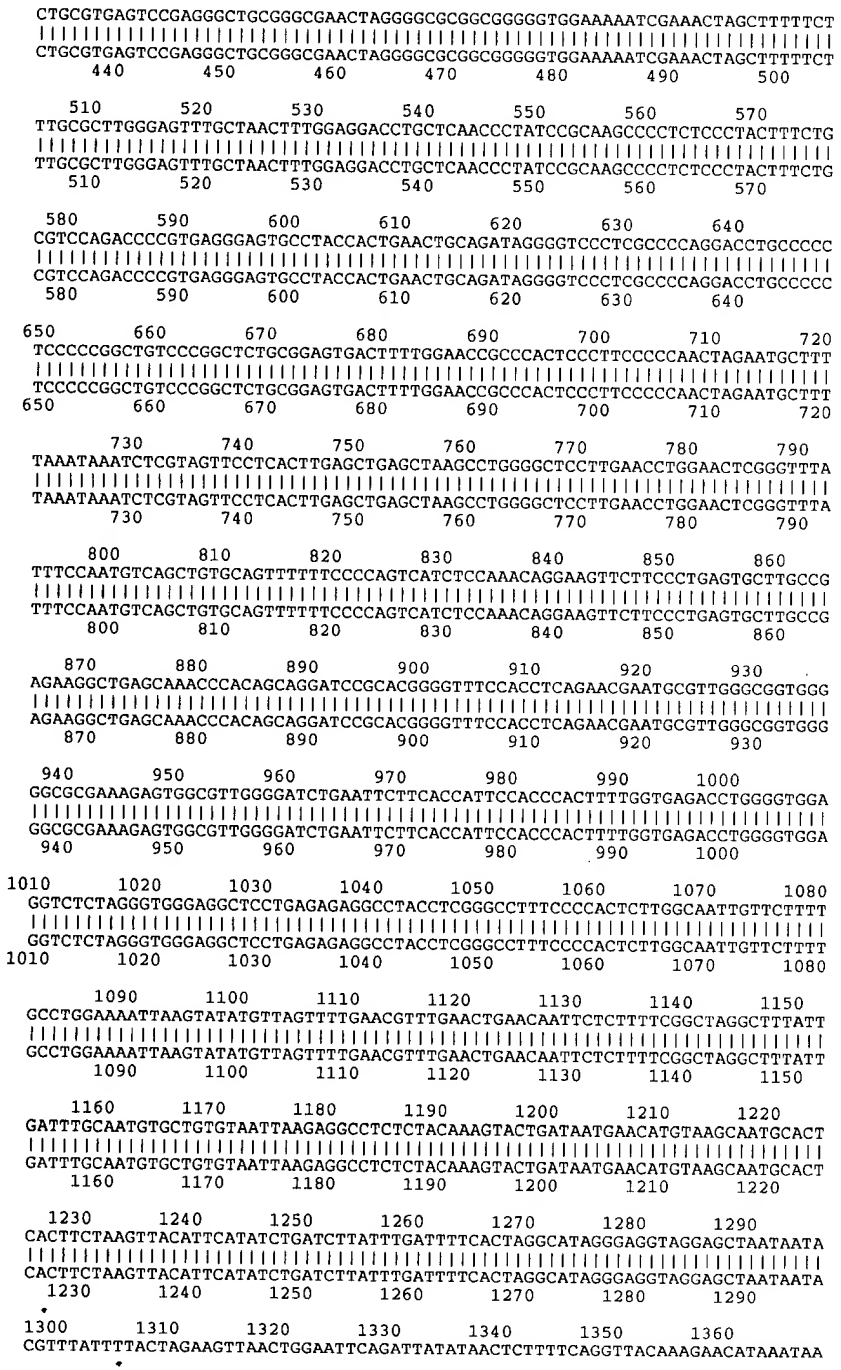
```
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-7


  Query Match            100.0%;  Score 301;  DB 4;  Length 10825;
  Best Local Similarity  100.0%;  Pred. No. 6.6e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    4690 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    4750 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    4810 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 4869

Qy     301 C 301
           |
Db    4870 C 4870


RESULT  13
US-09-277-457-27
; Sequence 27, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 12146
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-277-457-27


  Query Match            100.0%;  Score 301;  DB 4;  Length 12146;
  Best Local Similarity  100.0%;  Pred. No. 6.9e-86;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    5460 ggcacggaatccctggttggagtttcagaggtggctgaggctgtgtgcctctccaaattc 5519

Qy      61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    5520 tgggaagggactttctcaatcctagagtctctaccttataattgagatgtatgagacagc 5579

Qy     121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    5580 cacaagtcatgggtttaatttcttttctccatgcatatggctcaaagggaagtgtctatg 5639

Qy     181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    5640 gcccttgcttttttatttaaccaataatcttttgtatatttatacctgttaaaaattcaga 5699

Qy     241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    5700 aatgtcaaggccgggcacggtggctcaccccTgtaatcccagcactttgggaggccgagg 5759

Qy     301 C 301
           |
Db    5760 c 5760
```

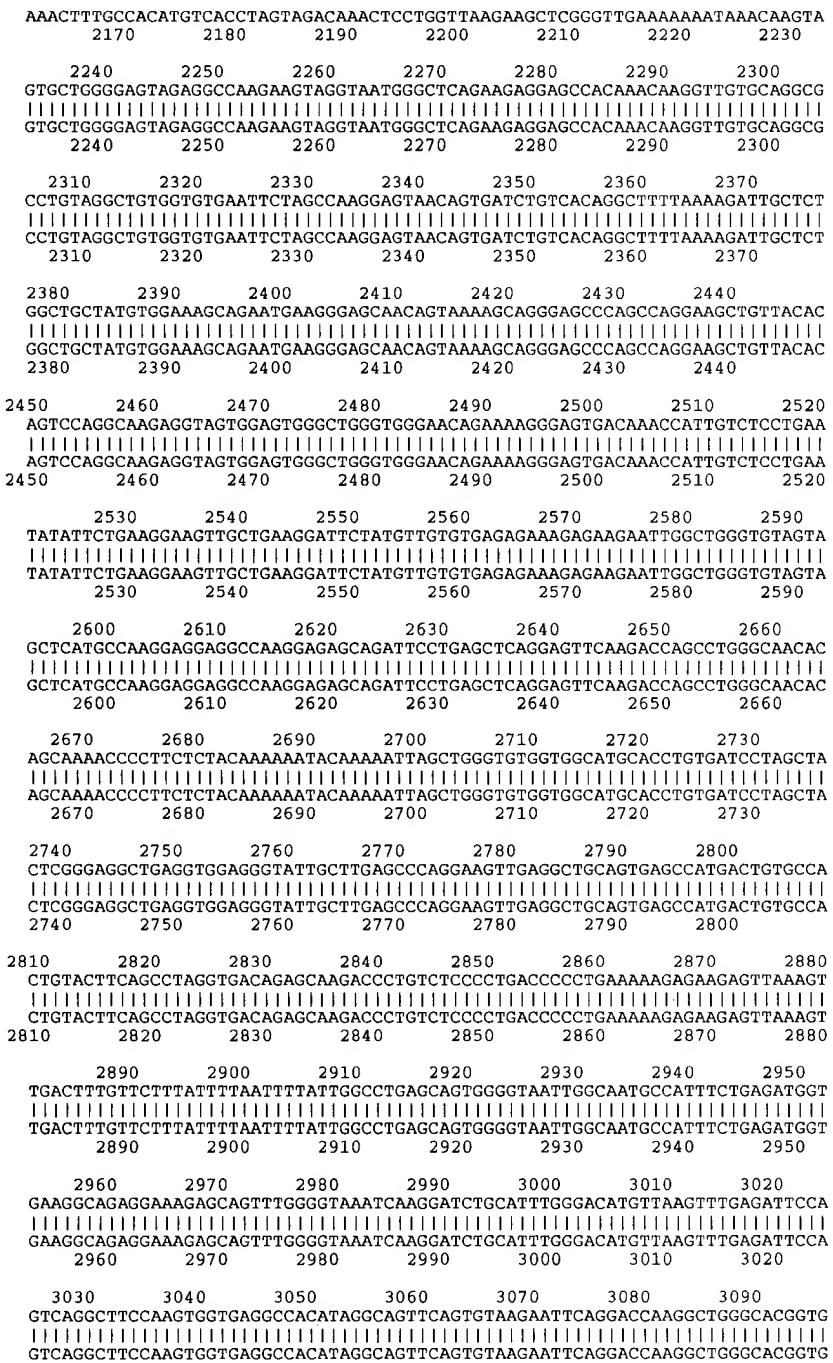
```
RESULT  14
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-20


  Query Match             100.0%;  Score 301;  DB 2;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-85;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196873 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 196932

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196933 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 196992

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196993 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 197052

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197053 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 197112

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197113 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 197172

Qy    301 C 301
          |
Db 197173 C 197173


RESULT  15
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-21


  Query Match             100.0%;  Score 301;  DB 2;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-85;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196873 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 196932

Qy     61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196933 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 196992

Qy    121 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
```

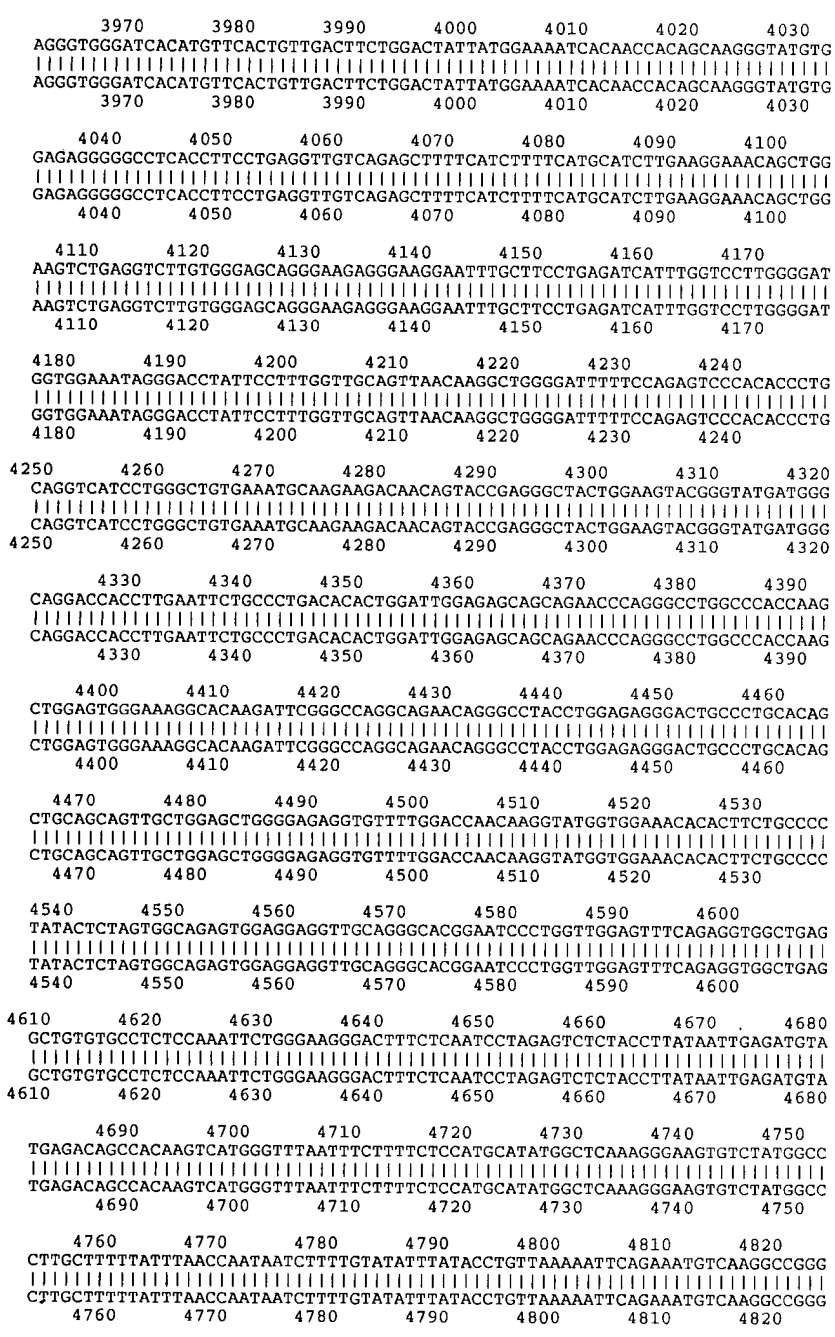
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196993 CACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 197052

Qy    181 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197053 GCCCTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 197112

Qy    241 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197113 AATGTCAAGGCCGGGCACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGG 197172

Qy    301 C 301
          |
Db 197173 C 197173
```

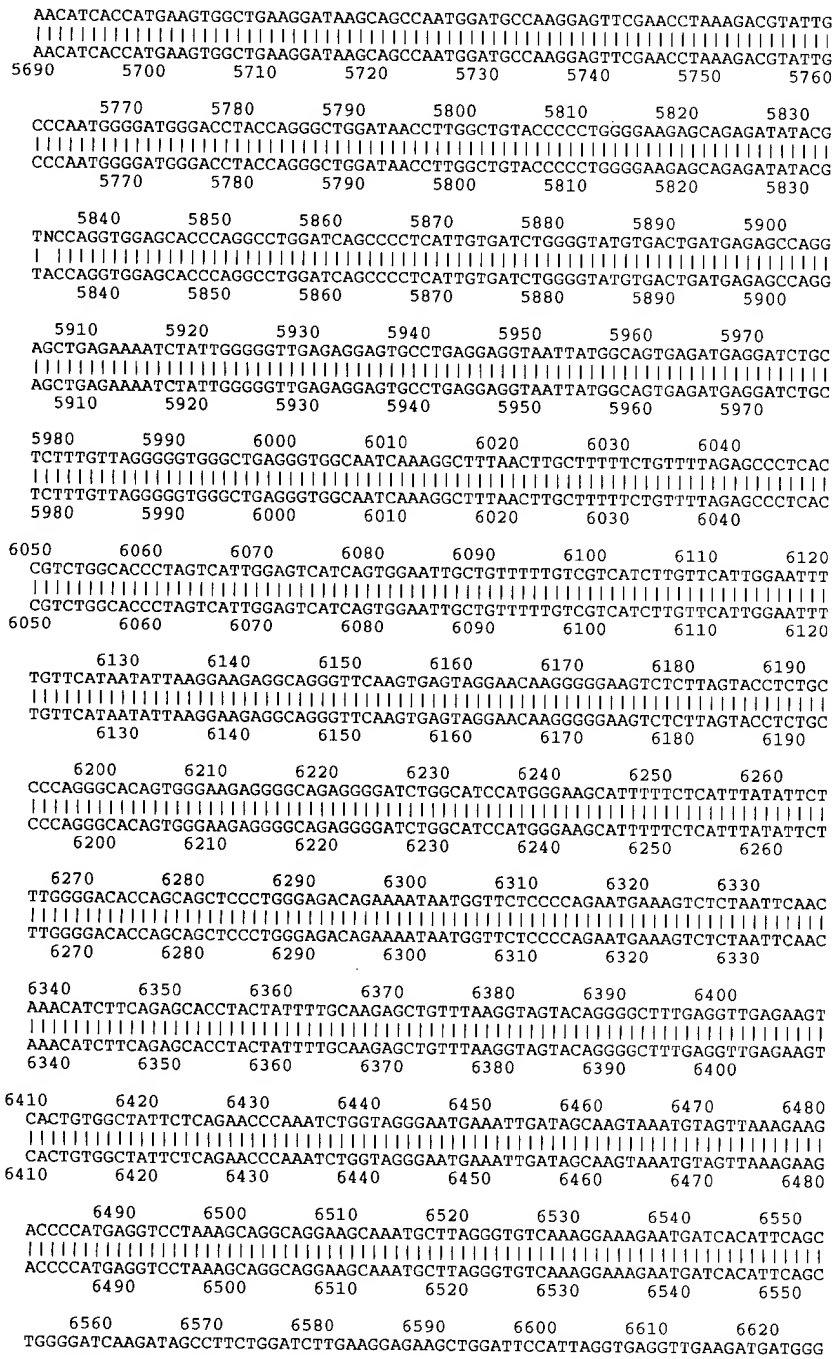
Search completed: June 19, 2002, 10:58:07
Job time: 8225 sec

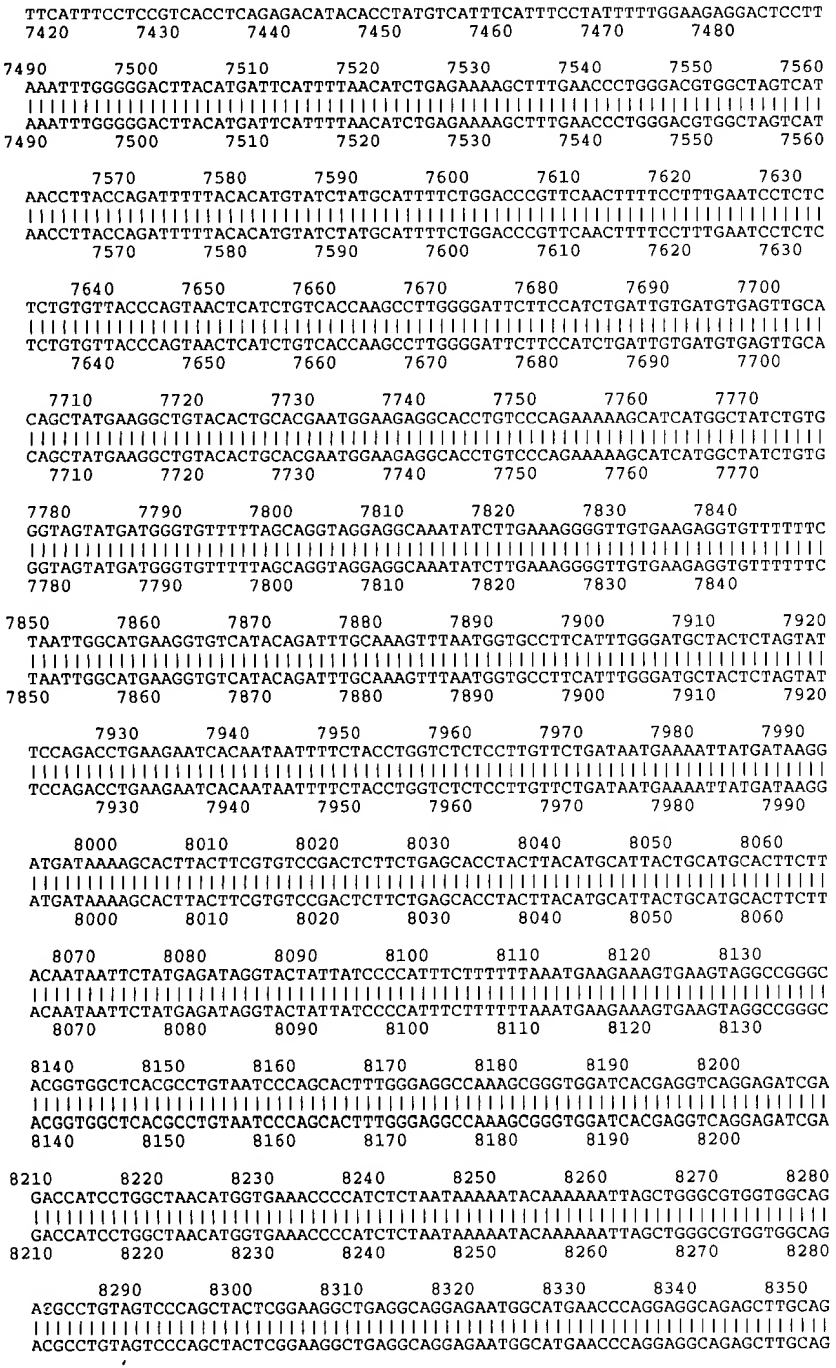
```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-497-957-1.res made by jdelaval on Wed 19 Jun 102 8:27:38-PDT.


Query sequence being compared:US-09-497-957-1 (1-10825)
Number of sequences searched:                  70
Number of scores above cutoff:                 70

      Results of the initial comparison of US-09-497-957-1 (1-10825) with:
        File : US09497957.seq

   100-
      -
N     *
U   50-
M     -
B     -
E     -
R     -
      -
O     -
F   10-
      -
S     -
E    5-  *                                                          *
Q     -
U     -
E     -
N     -
C     -
E     -   *
S     0-------------------------------------------------------------------
      |     |    | |     |     ||     |     |     |     ||     |
SCORE 0   1203   2406   3608   4811   6014   7217   8419   9622| 10825
STDEV           1            2            3            4


                        PARAMETERS

Similarity matrix      Unitary     K-tuple                    4
Mismatch penalty             1     Joining penalty           30
Gap penalty               1.00     Window size               32
Gap size penalty          0.33
Cutoff score                 0
Randomization group          0

                     SEARCH STATISTICS

Scores:                  Mean       Median     Standard Deviation
                          686           19     2519.87

Times:                    CPU                   Total Elapsed
                   00:00:02.02                  00:00:02.00

Number of residues:                51359
Number of sequences searched:         70
Number of scores above cutoff:        70

The scores below are sorted by initial score.
Significance is calculated based on initial score.

4 100% similar sequences to the query sequence were found:

                                                          Init. Opt.
Sequence Name     Description                    Length Score Score  Sig. Frame
-------------------------------------------------------------------------------
   1. US-09-497-957-7 Sequence 7, Application US  10825 10825 10825   4.02   0
   2. US-09-497-957-5 Sequence 5, Application US  10825 10825 10825   4.02   0
   3. US-09-497-957-3 Sequence 3, Application US  10825 10825 10825   4.02   0
   4. US-09-497-957-1 Sequence 1, Application US  10825 10825 10825   4.02   0

The list of other best scores is:

                                                          Init. Opt.
Sequence Name     Description                    Length Score Score  Sig. Frame
-------------------------------------------------------------------------------
   5. US-09-497-957-9 Sequence 9, Application US   1440   677   786  -0.00   0
   6. US-09-497-957-11 Sequence 11, Application    1440   676   785  -0.00   0
   7. US-09-497-957-10 Sequence 10, Application    1440   676   786  -0.00   0
   8. US-09-497-957-12 Sequence 12, Application    1440   675   785  -0.00   0
   9. US-09-497-957-21 Sequence 21, Application     517   516   516  -0.07   0
  10. US-09-497-957-20 Sequence 20, Application     517   516   516  -0.07   0
  11. US-09-497-957-42 Sequence 42, Application      40    39    39  -0.26   0
  12. US-09-497-957-41 Sequence 41, Application      40    39    39  -0.26   0
  13. US-09-497-957-30 Sequence 30, Application      39    38    38  -0.26   0
  14. US-09-497-957-29 Sequence 29, Application      39    38    38  -0.26   0
  15. US-09-497-957-43 Sequence 43, Application      32    31    31  -0.26   0
  16. US-09-497-957-44 Sequence 44, Application      32    30    30  -0.26   0
  17. US-09-497-957-31 Sequence 31, Application      22    22    22  -0.26   0
  18. US-09-497-957-24 Sequence 24, Application      20    20    20  -0.26   0
  19. US-09-497-957-18 Sequence 18, Application      20    20    20  -0.26   0
  20. US-09-497-957-13 Sequence 13, Application      20    20    20  -0.26   0


1. US-09-497-957-1 (1-10825)
   US-09-497-957-7 Sequence 7, Application US/09497957

Initial Score     =  10825  Optimized Score  =  10825  Significance =  4.02
Residue Identity =    99%  Matches          =  10823  Mismatches   =     2
Gaps              =      0  Conservative Substitutions               =     0

    X       10        20        30        40        50        60        70
    TCTAAGGTTGAGATAAAATTTTTAAATGTATGATTGAATTTTGAAAATCATAAATATTTAAATATCTAAAGT
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    TCTAAGGTTGAGATAAAATTTTTAAATGTATGATTGAATTTTGAAAATCATAAATATTTAAATATCTAAAGT
    X       10        20        30        40        50        60        70

          80        90       100       110       120       130       140
    TCAGATCAGAACATTGCGAAGCTACTTTCCCCAATCAACAACACCCCTTCAGGATTTAAAAACCAAGGGGGA
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    TCAGATCAGAACATTGCGAAGCTACTTTCCCCAATCAACAACACCCCTTCAGGATTTAAAAACCAAGGGGGA
          80        90       100       110       120       130       140

       150       160       170       180       190       200       210
    CACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGAACTAAAGTTCTGAAAGA
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    CACTGGATCACCTAGTGTTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGAACTAAAGTTCTGAAAGA
       150       160       170       180       190       200       210

     220       230       240       250       260       270       280
    CCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTG
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    CCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAGCCTAGGCAATAGCTGTAGGGTGACTTCTG
     220       230       240       250       260       270       280

290       300       310       320       330       340       350       360
    GAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAA
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    GAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAA
290       300       310       320       330       340       350       360

          370       380       390       400       410       420       430
    ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTG
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGCTTG
          370       380       390       400       410       420       430

       440       450       460       470       480       490       500
```

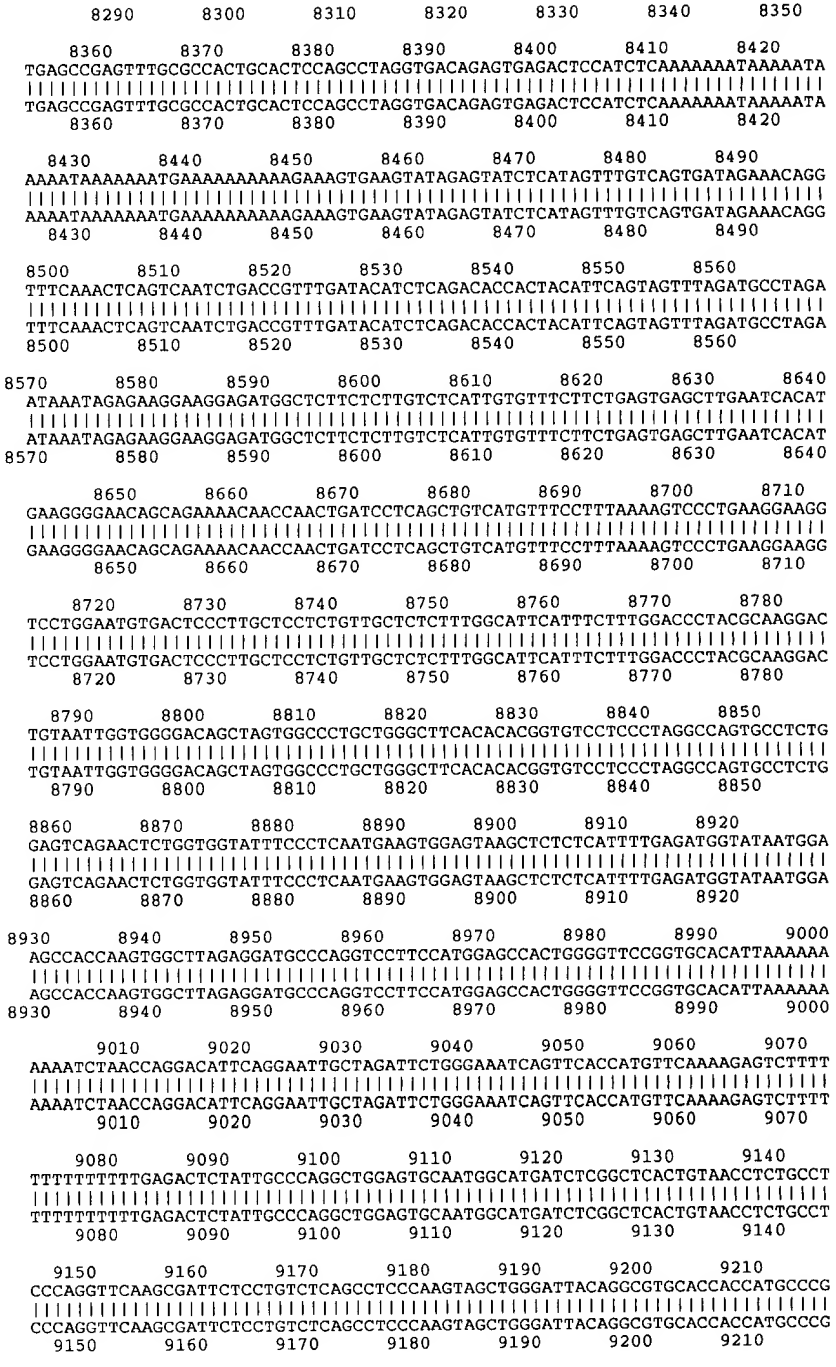

CTGCGTGAGTCCGAGGGCTGCGGGCGAACTAGGGGCGCGGCGGGGGGTGGAAAAATCGAAACTAGCTTTTTCT
CTGCGTGAGTCCGAGGGCTGCGGGCGAACTAGGGGCGCGGCGGGGGGTGGAAAAATCGAAACTAGCTTTTTCT
440 450 460 470 480 490 500
510 520 530 540 550 560 570
TTGCGCTTGGGAGTTTGCTAACTTTGGAGGACCTGCTCAACCCTATCCGCAAGCCCCTCTCCCTACTTTCTG
TTGCGCTTGGGAGTTTGCTAACTTTGGAGGACCTGCTCAACCCTATCCGCAAGCCCCTCTCCCTACTTTCTG
510 520 530 540 550 560 570
580 590 600 610 620 630 640
CGTCCAGACCCCGTGAGGGAGTGCCTACCACTGAACTGCAGATAGGGGTCCCTCGCCCCAGGACCTGCCCCC
CGTCCAGACCCCGTGAGGGAGTGCCTACCACTGAACTGCAGATAGGGGTCCCTCGCCCCAGGACCTGCCCCC
580 590 600 610 620 630 640
650 660 670 680 690 700 710 720
TCCCCCGGCTGTCCCGGCTCTGCGGAGTGACTTTTGGAACCGCCCACTCCCTTCCCCCAACTAGAATGCTTT
TCCCCCGGCTGTCCCGGCTCTGCGGAGTGACTTTTGGAACCGCCCACTCCCTTCCCCCAACTAGAATGCTTT
650 660 670 680 690 700 710 720
730 740 750 760 770 780 790
TAAATAAATCTCGTAGTTCCTCACTTGAGCTGAGCTAAGCCTGGGGCTCCTTGAACCTGGAACTCGGGTTTA
TAAATAAATCTCGTAGTTCCTCACTTGAGCTGAGCTAAGCCTGGGGCTCCTTGAACCTGGAACTCGGGTTTA
730 740 750 760 770 780 790
800 810 820 830 840 850 860
TTTCCAATGTCAGCTGTGCAGTTTTTTCCCCAGTCATCTCCAAACAGGAAGTTCTTCCCTGAGTGCTTGCCG
TTTCCAATGTCAGCTGTGCAGTTTTTTCCCCAGTCATCTCCAAACAGGAAGTTCTTCCCTGAGTGCTTGCCG
800 810 820 830 840 850 860
870 880 890 900 910 920 930
AGAAGGCTGAGCAAACCCACAGCAGGATCCGCACGGGGTTTCCACCTCAGAACGAATGCGTTGGGCGGTGGG
AGAAGGCTGAGCAAACCCACAGCAGGATCCGCACGGGGTTTCCACCTCAGAACGAATGCGTTGGGCGGTGGG
870 880 890 900 910 920 930
940 950 960 970 980 990 1000
GGCGCGAAAGAGTGGCGTTGGGGATCTGAATTCTTCACCATTCCACCCACTTTTGGTGAGACCTGGGGTGGA
GGCGCGAAAGAGTGGCGTTGGGGATCTGAATTCTTCACCATTCCACCCACTTTTGGTGAGACCTGGGGTGGA
940 950 960 970 980 990 1000
1010 1020 1030 1040 1050 1060 1070 1080
GGTCTCTAGGGTGGGAGGCTCCTGAGAGAGGCCTACCTCGGGCCTTTCCCCACTCTTGGCAATTGTTCTTTT
GGTCTCTAGGGTGGGAGGCTCCTGAGAGAGGCCTACCTCGGGCCTTTCCCCACTCTTGGCAATTGTTCTTTT
1010 1020 1030 1040 1050 1060 1070 1080
1090 1100 1110 1120 1130 1140 1150
GCCTGGAAAATTAAGTATATGTTAGTTTTGAACGTTTGAACTGAACAATTCTCTTTTCGGCTAGGCTTTATT
GCCTGGAAAATTAAGTATATGTTAGTTTTGAACGTTTGAACTGAACAATTCTCTTTTCGGCTAGGCTTTATT
1090 1100 1110 1120 1130 1140 1150
1160 1170 1180 1190 1200 1210 1220
GATTTGCAATGTGCTGTGTAATTAAGAGGCCTCTCTACAAAGTACTGATAATGAACATGTAAGCAATGCACT
GATTTGCAATGTGCTGTGTAATTAAGAGGCCTCTCTACAAAGTACTGATAATGAACATGTAAGCAATGCACT
1160 1170 1180 1190 1200 1210 1220
1230 1240 1250 1260 1270 1280 1290
CACTTCTAAGTTACATTCATATCTGATCTTATTTGATTTTCACTAGGCATAGGGAGGTAGGAGCTAATAATA
CACTTCTAAGTTACATTCATATCTGATCTTATTTGATTTTCACTAGGCATAGGGAGGTAGGAGCTAATAATA
1230 1240 1250 1260 1270 1280 1290
1300 1310 1320 1330 1340 1350 1360
CGTTTATTTTACTAGAAGTTAACTGGAATTCAGATTATATAACTCTTTTCAGGTTACAAAGAACATAAATAA
CGTTTATTTTACTAGAAGTTAACTGGAATTCAGATTATATAACTCTTTTCAGGTTACAAAGAACATAAATAA
1300 1310 1320 1330 1340 1350 1360
1370 1380 1390 1400 1410 1420 1430 1440
TCTGGTTTTCTGATGTTATTTCAAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTGATTTTGCCCTGTAG
TCTGGTTTTCTGATGTTATTTCAAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTGATTTTGCCCTGTAG
1370 1380 1390 1400 1410 1420 1430 1440
1450 1460 1470 1480 1490 1500 1510
TGTAGCACAGTGTTCTGTGGGTCACACGCCGGCCTCAGCACAGCACTTTGAGTTTTGGTACTACGTGTATCC
TGTAGCACAGTGTTCTGTGGGTCACACGCCGGCCTCAGCACAGCACTTTGAGTTTTGGTACTACGTGTATCC
1450 1460 1470 1480 1490 1500 1510
1520 1530 1540 1550 1560 1570 1580
ACATTTTACACATGACAAGAATGAGGCATGGCACGGCCTGCTTCCTGGCAAATTTATTCAATGGTACACTGG
ACATTTTACACATGACAAGAATGAGGCATGGCACGGCCTGCTTCCTGGCAAATTTATTCAATGGTACACTGG
1520 1530 1540 1550 1560 1570 1580
1590 1600 1610 1620 1630 1640 1650
GCTTTGGTGGCAGAGCTCATGTCTCCACTTCATAGCTATGATTCTTAAACATCACACTGCATTAGAGGTTGA
GCTTTGGTGGCAGAGCTCATGTCTCCACTTCATAGCTATGATTCTTAAACATCACACTGCATTAGAGGTTGA
1590 1600 1610 1620 1630 1640 1650
1660 1670 1680 1690 1700 1710 1720
ATAATAAAATTTCATGTTGAGCAGAAATATTCATTGTTTACAAGTGTAAATGAGTCCCAGCCATGTGTTGCA
ATAATAAAATTTCATGTTGAGCAGAAATATTCATTGTTTACAAGTGTAAATGAGTCCCAGCCATGTGTTGCA
1660 1670 1680 1690 1700 1710 1720
1730 1740 1750 1760 1770 1780 1790 1800
CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
1730 1740 1750 1760 1770 1780 1790 1800
1810 1820 1830 1840 1850 1860 1870
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
1810 1820 1830 1840 1850 1860 1870
1880 1890 1900 1910 1920 1930 1940
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
1880 1890 1900 1910 1920 1930 1940
1950 1960 1970 1980 1990 2000 2010
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
1950 1960 1970 1980 1990 2000 2010
2020 2030 2040 2050 2060 2070 2080
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
2020 2030 2040 2050 2060 2070 2080
2090 2100 2110 2120 2130 2140 2150 2160
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAAATGTCTCT
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAAATGTCTCT
2090 2100 2110 2120 2130 2140 2150 2160
2170 2180 2190 2200 2210 2220 2230
AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA

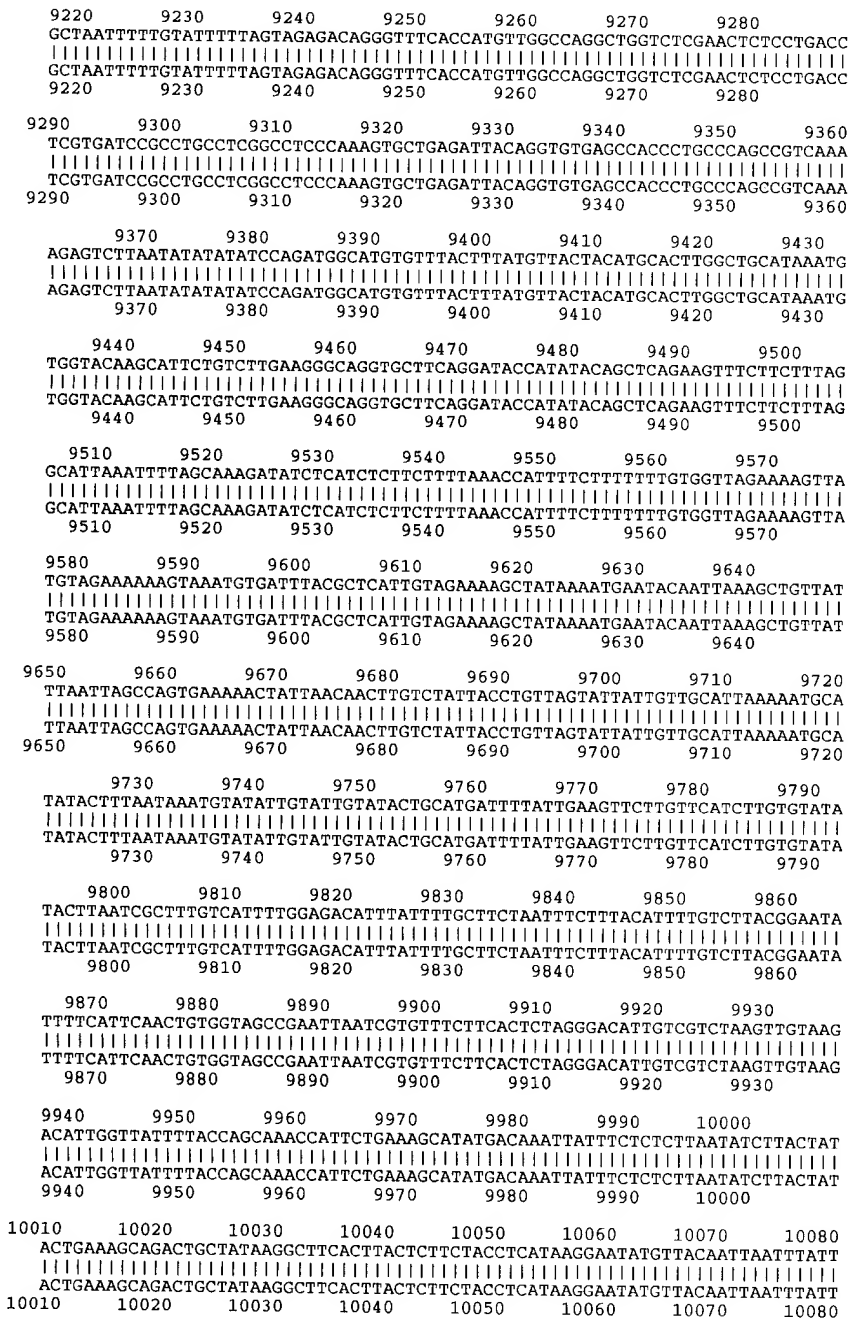

AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA
2170 2180 2190 2200 2210 2220 2230
2240 2250 2260 2270 2280 2290 2300
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
2240 2250 2260 2270 2280 2290 2300
2310 2320 2330 2340 2350 2360 2370
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
2310 2320 2330 2340 2350 2360 2370
2380 2390 2400 2410 2420 2430 2440
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
2380 2390 2400 2410 2420 2430 2440
2450 2460 2470 2480 2490 2500 2510 2520
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
2450 2460 2470 2480 2490 2500 2510 2520
2530 2540 2550 2560 2570 2580 2590
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
2530 2540 2550 2560 2570 2580 2590
2600 2610 2620 2630 2640 2650 2660
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
2600 2610 2620 2630 2640 2650 2660
2670 2680 2690 2700 2710 2720 2730
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
2670 2680 2690 2700 2710 2720 2730
2740 2750 2760 2770 2780 2790 2800
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
2740 2750 2760 2770 2780 2790 2800
2810 2820 2830 2840 2850 2860 2870 2880
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
2810 2820 2830 2840 2850 2860 2870 2880
2890 2900 2910 2920 2930 2940 2950
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
2890 2900 2910 2920 2930 2940 2950
2960 2970 2980 2990 3000 3010 3020
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
2960 2970 2980 2990 3000 3010 3020
3030 3040 3050 3060 3070 3080 3090
GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
3030 3040 3050 3060 3070 3080 3090
3100 3110 3120 3130 3140 3150 3160
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
3100 3110 3120 3130 3140 3150 3160
3170 3180 3190 3200 3210 3220 3230 3240
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
3170 3180 3190 3200 3210 3220 3230 3240
3250 3260 3270 3280 3290 3300 3310
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
3250 3260 3270 3280 3290 3300 3310
3320 3330 3340 3350 3360 3370 3380
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
3320 3330 3340 3350 3360 3370 3380
3390 3400 3410 3420 3430 3440 3450
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
3390 3400 3410 3420 3430 3440 3450
3460 3470 3480 3490 3500 3510 3520
TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA
TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA
3460 3470 3480 3490 3500 3510 3520
3530 3540 3550 3560 3570 3580 3590 3600
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
3530 3540 3550 3560 3570 3580 3590 3600
3610 3620 3630 3640 3650 3660 3670
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
3610 3620 3630 3640 3650 3660 3670
3680 3690 3700 3710 3720 3730 3740
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
3680 3690 3700 3710 3720 3730 3740
3750 3760 3770 3780 3790 3800 3810
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
3750 3760 3770 3780 3790 3800 3810
3820 3830 3840 3850 3860 3870 3880
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATNATGAGAGTCGCCGTGT
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGT
3820 3830 3840 3850 3860 3870 3880
3890 3900 3910 3920 3930 3940 3950 3960
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
3890 3900 3910 3920 3930 3940 3950 3960

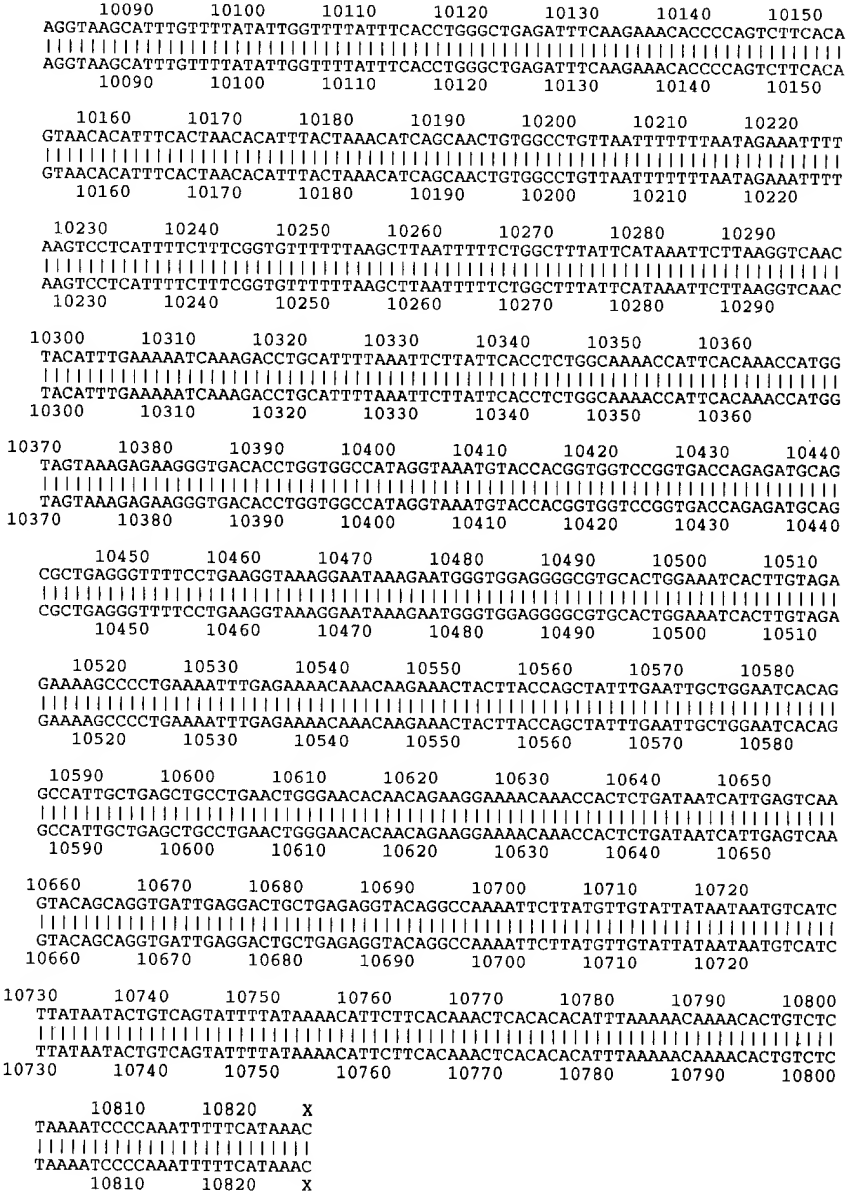

3970 3980 3990 4000 4010 4020 4030
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
3970 3980 3990 4000 4010 4020 4030
4040 4050 4060 4070 4080 4090 4100
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
4040 4050 4060 4070 4080 4090 4100
4110 4120 4130 4140 4150 4160 4170
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
4110 4120 4130 4140 4150 4160 4170
4180 4190 4200 4210 4220 4230 4240
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
4180 4190 4200 4210 4220 4230 4240
4250 4260 4270 4280 4290 4300 4310 4320
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
4250 4260 4270 4280 4290 4300 4310 4320
4330 4340 4350 4360 4370 4380 4390
CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
4330 4340 4350 4360 4370 4380 4390
4400 4410 4420 4430 4440 4450 4460
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
4400 4410 4420 4430 4440 4450 4460
4470 4480 4490 4500 4510 4520 4530
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
4470 4480 4490 4500 4510 4520 4530
4540 4550 4560 4570 4580 4590 4600
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
4540 4550 4560 4570 4580 4590 4600
4610 4620 4630 4640 4650 4660 4670 4680
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
4610 4620 4630 4640 4650 4660 4670 4680
4690 4700 4710 4720 4730 4740 4750
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
4690 4700 4710 4720 4730 4740 4750
4760 4770 4780 4790 4800 4810 4820
CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGG
CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGG
4760 4770 4780 4790 4800 4810 4820
4830 4840 4850 4860 4870 4880 4890
CACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGTCACAAGGTCAGGAGTTTGA
CACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGTCACAAGGTCAGGAGTTTGA
4830 4840 4850 4860 4870 4880 4890
4900 4910 4920 4930 4940 4950 4960
GACCAGCCTGACCAACATGGTGAAACCCGTCTCTAAAAAAATACAAAAATTAGCTGGTCACAGTCATGCGCA
GACCAGCCTGACCAACATGGTGAAACCCGTCTCTAAAAAAATACAAAAATTAGCTGGTCACAGTCATGCGCA
4900 4910 4920 4930 4940 4950 4960
4970 4980 4990 5000 5010 5020 5030 5040
CCTGTAGTCCCAGCTAATTGGAAGGCTGAGGCAGGAGCATCGCTTGAACCTGGGAAGCGGAAGTTGCACTGA
CCTGTAGTCCCAGCTAATTGGAAGGCTGAGGCAGGAGCATCGCTTGAACCTGGGAAGCGGAAGTTGCACTGA
4970 4980 4990 5000 5010 5020 5030 5040
5050 5060 5070 5080 5090 5100 5110
GCCAAGATCGCGCCACTGCACTCCAGCCTAGGCAGCAGAGTGAGACTCCATCTTAAAAAAAAAAAAAAAAAA
GCCAAGATCGCGCCACTGCACTCCAGCCTAGGCAGCAGAGTGAGACTCCATCTTAAAAAAAAAAAAAAAAAA
5050 5060 5070 5080 5090 5100 5110
5120 5130 5140 5150 5160 5170 5180
AAAAAGAGAATTCAGAGATCTCAGCTATCATATGAATACCAGGACAAAATATCAAGTGAGGCCACTTATCAG
AAAAAGAGAATTCAGAGATCTCAGCTATCATATGAATACCAGGACAAAATATCAAGTGAGGCCACTTATCAG
5120 5130 5140 5150 5160 5170 5180
5190 5200 5210 5220 5230 5240 5250
AGTAGAAGAATCCTTTAGGTTAAAAGTTTCTTTCATAGAACATAGCAATAATCACTGAAGCTACCTATCTTA
AGTAGAAGAATCCTTTAGGTTAAAAGTTTCTTTCATAGAACATAGCAATAATCACTGAAGCTACCTATCTTA
5190 5200 5210 5220 5230 5240 5250
5260 5270 5280 5290 5300 5310 5320
CAAGTCCGCTTCTTATAACAATGCCTCCTAGGTTGACCCAGGTGAAACTGACCATCTGTATTCAATCATTTT
CAAGTCCGCTTCTTATAACAATGCCTCCTAGGTTGACCCAGGTGAAACTGACCATCTGTATTCAATCATTTT
5260 5270 5280 5290 5300 5310 5320
5330 5340 5350 5360 5370 5380 5390 5400
CAATGCACATAAAGGGCAATTTTATCTATCAGAACAAAGAACATGGGTAACAGATATGTATATTTACATGTG
CAATGCACATAAAGGGCAATTTTATCTATCAGAACAAAGAACATGGGTAACAGATATGTATATTTACATGTG
5330 5340 5350 5360 5370 5380 5390 5400
5410 5420 5430 5440 5450 5460 5470
AGGAGAACAAGCTGATCTGACTGCTCTCCAAGTGACACTGTGTTAGAGTCCAATCTTAGGACACAAAATGGT
AGGAGAACAAGCTGATCTGACTGCTCTCCAAGTGACACTGTGTTAGAGTCCAATCTTAGGACACAAAATGGT
5410 5420 5430 5440 5450 5460 5470
5480 5490 5500 5510 5520 5530 5540
GTCTCTCCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTT
GTCTCTCCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTT
5480 5490 5500 5510 5520 5530 5540
5550 5560 5570 5580 5590 5600 5610
CCAGTCTTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTG
CCAGTCTTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTG
5550 5560 5570 5580 5590 5600 5610
5620 5630 5640 5650 5660 5670 5680
GTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG
GTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG
5620 5630 5640 5650 5660 5670 5680
5690 5700 5710 5720 5730 5740 5750 5760


AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTG
AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTG
5690 5700 5710 5720 5730 5740 5750 5760
5770 5780 5790 5800 5810 5820 5830
CCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACG
CCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACG
5770 5780 5790 5800 5810 5820 5830
5840 5850 5860 5870 5880 5890 5900
TNCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGG
TACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGG
5840 5850 5860 5870 5880 5890 5900
5910 5920 5930 5940 5950 5960 5970
AGCTGAGAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGC
AGCTGAGAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGC
5910 5920 5930 5940 5950 5960 5970
5980 5990 6000 6010 6020 6030 6040
TCTTTGTTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCAC
TCTTTGTTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCAC
5980 5990 6000 6010 6020 6030 6040
6050 6060 6070 6080 6090 6100 6110 6120
CGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTGTTCATTGGAATTT
CGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTGTTCATTGGAATTT
6050 6060 6070 6080 6090 6100 6110 6120
6130 6140 6150 6160 6170 6180 6190
TGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGTGAGTAGGAACAAGGGGGAAGTCTCTTAGTACCTCTGC
TGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGTGAGTAGGAACAAGGGGGAAGTCTCTTAGTACCTCTGC
6130 6140 6150 6160 6170 6180 6190
6200 6210 6220 6230 6240 6250 6260
CCCAGGGCACAGTGGGAAGAGGGGCAGAGGGGATCTGGCATCCATGGGAAGCATTTTTCTCATTTATATTCT
CCCAGGGCACAGTGGGAAGAGGGGCAGAGGGGATCTGGCATCCATGGGAAGCATTTTTCTCATTTATATTCT
6200 6210 6220 6230 6240 6250 6260
6270 6280 6290 6300 6310 6320 6330
TTGGGGACACCAGCAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCCAGAATGAAAGTCTCTAATTCAAC
TTGGGGACACCAGCAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCCAGAATGAAAGTCTCTAATTCAAC
6270 6280 6290 6300 6310 6320 6330
6340 6350 6360 6370 6380 6390 6400
AAACATCTTCAGAGCACCTACTATTTTGCAAGAGCTGTTTAAGGTAGTACAGGGGCTTTGAGGTTGAGAAGT
AAACATCTTCAGAGCACCTACTATTTTGCAAGAGCTGTTTAAGGTAGTACAGGGGCTTTGAGGTTGAGAAGT
6340 6350 6360 6370 6380 6390 6400
6410 6420 6430 6440 6450 6460 6470 6480
CACTGTGGCTATTCTCAGAACCCAAATCTGGTAGGGAATGAAATTGATAGCAAGTAAATGTAGTTAAAGAAG
CACTGTGGCTATTCTCAGAACCCAAATCTGGTAGGGAATGAAATTGATAGCAAGTAAATGTAGTTAAAGAAG
6410 6420 6430 6440 6450 6460 6470 6480
6490 6500 6510 6520 6530 6540 6550
ACCCCATGAGGTCCTAAAGCAGGCAGGAAGCAAATGCTTAGGGTGTCAAAGGAAAGAATGATCACATTCAGC
ACCCCATGAGGTCCTAAAGCAGGCAGGAAGCAAATGCTTAGGGTGTCAAAGGAAAGAATGATCACATTCAGC
6490 6500 6510 6520 6530 6540 6550
6560 6570 6580 6590 6600 6610 6620
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
6560 6570 6580 6590 6600 6610 6620
6630 6640 6650 6660 6670 6680 6690
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
6630 6640 6650 6660 6670 6680 6690
6700 6710 6720 6730 6740 6750 6760
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
6700 6710 6720 6730 6740 6750 6760
6770 6780 6790 6800 6810 6820 6830 6840
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
6770 6780 6790 6800 6810 6820 6830 6840
6850 6860 6870 6880 6890 6900 6910
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
6850 6860 6870 6880 6890 6900 6910
6920 6930 6940 6950 6960 6970 6980
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT
6920 6930 6940 6950 6960 6970 6980
6990 7000 7010 7020 7030 7040 7050
TGAAATCCAATAGTGCCCAGGTCTAAATTGAGATGGGTGAATGAGGAAAATAAGGAAGAGAGAAGAGGCAAG
TGAAATCCAATAGTGCCCAGGTCTAAATTGAGATGGGTGAATGAGGAAAATAAGGAAGAGAGAAGAGGCAAG
6990 7000 7010 7020 7030 7040 7050
7060 7070 7080 7090 7100 7110 7120
ATGGTGCCTAGGTTTGTGATGCCTCTTTCCTGGGTCTCTTGTCTCCACAGGAGGAGCCATGGGGCACTACGT
ATGGTGCCTAGGTTTGTGATGCCTCTTTCCTGGGTCTCTTGTCTCCACAGGAGGAGCCATGGGGCACTACGT
7060 7070 7080 7090 7100 7110 7120
7130 7140 7150 7160 7170 7180 7190 7200
CTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA
CTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA
7130 7140 7150 7160 7170 7180 7190 7200
7210 7220 7230 7240 7250 7260 7270
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACT
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACT
7210 7220 7230 7240 7250 7260 7270
7280 7290 7300 7310 7320 7330 7340
CCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGC
CCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGC
7280 7290 7300 7310 7320 7330 7340
7350 7360 7370 7380 7390 7400 7410
CGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCC
CGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCC
7350 7360 7370 7380 7390 7400 7410
7420 7430 7440 7450 7460 7470 7480
TTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTT


TTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTT
7420 7430 7440 7450 7460 7470 7480
7490 7500 7510 7520 7530 7540 7550 7560
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCAT
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCAT
7490 7500 7510 7520 7530 7540 7550 7560
7570 7580 7590 7600 7610 7620 7630
AACCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTC
AACCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTC
7570 7580 7590 7600 7610 7620 7630
7640 7650 7660 7670 7680 7690 7700
TCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCA
TCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCA
7640 7650 7660 7670 7680 7690 7700
7710 7720 7730 7740 7750 7760 7770
CAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTG
CAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTG
7710 7720 7730 7740 7750 7760 7770
7780 7790 7800 7810 7820 7830 7840
GGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGTGTTTTTTC
GGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGTGTTTTTTC
7780 7790 7800 7810 7820 7830 7840
7850 7860 7870 7880 7890 7900 7910 7920
TAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTAT
TAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTAT
7850 7860 7870 7880 7890 7900 7910 7920
7930 7940 7950 7960 7970 7980 7990
TCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGG
TCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGG
7930 7940 7950 7960 7970 7980 7990
8000 8010 8020 8030 8040 8050 8060
ATGATAAAAGCACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTT
ATGATAAAAGCACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTT
8000 8010 8020 8030 8040 8050 8060
8070 8080 8090 8100 8110 8120 8130
ACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGC
ACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGC
8070 8080 8090 8100 8110 8120 8130
8140 8150 8160 8170 8180 8190 8200
ACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGCGGGTGGATCACGAGGTCAGGAGATCGA
ACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGCGGGTGGATCACGAGGTCAGGAGATCGA
8140 8150 8160 8170 8180 8190 8200
8210 8220 8230 8240 8250 8260 8270 8280
GACCATCCTGGCTAACATGGTGAAACCCCATCTCTAATAAAAATACAAAAAATTAGCTGGGCGTGGTGGCAG
GACCATCCTGGCTAACATGGTGAAACCCCATCTCTAATAAAAATACAAAAAATTAGCTGGGCGTGGTGGCAG
8210 8220 8230 8240 8250 8260 8270 8280
8290 8300 8310 8320 8330 8340 8350
ACGCCTGTAGTCCCAGCTACTCGGAAGGCTGAGGCAGGAGAATGGCATGAACCCAGGAGGCAGAGCTTGCAG
ACGCCTGTAGTCCCAGCTACTCGGAAGGCTGAGGCAGGAGAATGGCATGAACCCAGGAGGCAGAGCTTGCAG

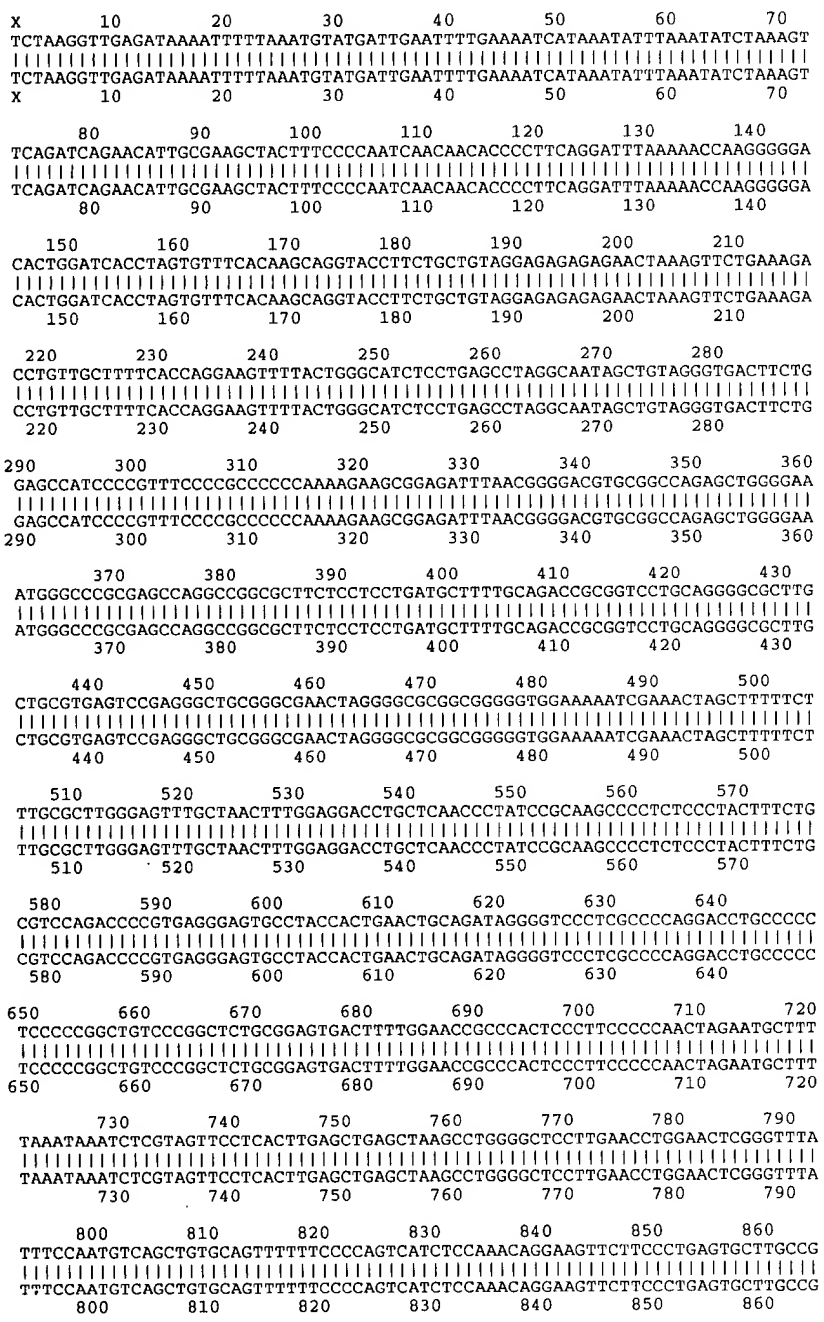

8290 8300 8310 8320 8330 8340 8350
8360 8370 8380 8390 8400 8410 8420
TGAGCCGAGTTTGCGCCACTGCACTCCAGCCTAGGTGACAGAGTGAGACTCCATCTCAAAAAAATAAAAATA
TGAGCCGAGTTTGCGCCACTGCACTCCAGCCTAGGTGACAGAGTGAGACTCCATCTCAAAAAAATAAAAATA
8360 8370 8380 8390 8400 8410 8420
8430 8440 8450 8460 8470 8480 8490
AAAATAAAAAAATGAAAAAAAAAAAGAAAGTGAAGTATAGAGTATCTCATAGTTTGTCAGTGATAGAAACAGG
AAAATAAAAAAATGAAAAAAAAAAAGAAAGTGAAGTATAGAGTATCTCATAGTTTGTCAGTGATAGAAACAGG
8430 8440 8450 8460 8470 8480 8490
8500 8510 8520 8530 8540 8550 8560
TTTCAAACTCAGTCAATCTGACCGTTTGATACATCTCAGACACCACTACATTCAGTAGTTTAGATGCCTAGA
TTTCAAACTCAGTCAATCTGACCGTTTGATACATCTCAGACACCACTACATTCAGTAGTTTAGATGCCTAGA
8500 8510 8520 8530 8540 8550 8560
8570 8580 8590 8600 8610 8620 8630 8640
ATAAATAGAGAAGGAAGGAGATGGCTCTTCTCTTGTCTCATTGTGTTTCTTCTGAGTGAGCTTGAATCACAT
ATAAATAGAGAAGGAAGGAGATGGCTCTTCTCTTGTCTCATTGTGTTTCTTCTGAGTGAGCTTGAATCACAT
8570 8580 8590 8600 8610 8620 8630 8640
8650 8660 8670 8680 8690 8700 8710
GAAGGGGAACAGCAGAAAACAACCAACTGATCCTCAGCTGTCATGTTTCCTTTAAAAGTCCCTGAAGGAAGG
GAAGGGGAACAGCAGAAAACAACCAACTGATCCTCAGCTGTCATGTTTCCTTTAAAAGTCCCTGAAGGAAGG
8650 8660 8670 8680 8690 8700 8710
8720 8730 8740 8750 8760 8770 8780
TCCTGGAATGTGACTCCCTTGCTCCTCTGTTGCTCTCTTTGGCATTCATTTCTTTGGACCCTACGCAAGGAC
TCCTGGAATGTGACTCCCTTGCTCCTCTGTTGCTCTCTTTGGCATTCATTTCTTTGGACCCTACGCAAGGAC
8720 8730 8740 8750 8760 8770 8780
8790 8800 8810 8820 8830 8840 8850
TGTAATTGGTGGGGACAGCTAGTGGCCCTGCTGGGCTTCACACACGGTGTCCTCCCTAGGCCAGTGCCTCTG
TGTAATTGGTGGGGACAGCTAGTGGCCCTGCTGGGCTTCACACACGGTGTCCTCCCTAGGCCAGTGCCTCTG
8790 8800 8810 8820 8830 8840 8850
8860 8870 8880 8890 8900 8910 8920
GAGTCAGAACTCTGGTGGTATTTCCCTCAATGAAGTGGAGTAAGCTCTCTCATTTTGAGATGGTATAATGGA
GAGTCAGAACTCTGGTGGTATTTCCCTCAATGAAGTGGAGTAAGCTCTCTCATTTTGAGATGGTATAATGGA
8860 8870 8880 8890 8900 8910 8920
8930 8940 8950 8960 8970 8980 8990 9000
AGCCACCAAGTGGCTTAGAGGATGCCCAGGTCCTTCCATGGAGCCACTGGGGTTCCGGTGCACATTAAAAAA
AGCCACCAAGTGGCTTAGAGGATGCCCAGGTCCTTCCATGGAGCCACTGGGGTTCCGGTGCACATTAAAAAA
8930 8940 8950 8960 8970 8980 8990 9000
9010 9020 9030 9040 9050 9060 9070
AAAATCTAACCAGGACATTCAGGAATTGCTAGATTCTGGGAAATCAGTTCACCATGTTCAAAAGAGTCTTTT
AAAATCTAACCAGGACATTCAGGAATTGCTAGATTCTGGGAAATCAGTTCACCATGTTCAAAAGAGTCTTTT
9010 9020 9030 9040 9050 9060 9070
9080 9090 9100 9110 9120 9130 9140
TTTTTTTTTTGAGACTCTATTGCCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGTAACCTCTGCCT
TTTTTTTTTTGAGACTCTATTGCCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGTAACCTCTGCCT
9080 9090 9100 9110 9120 9130 9140
9150 9160 9170 9180 9190 9200 9210
CCCAGGTTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCACCACCATGCCCG
CCCAGGTTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCACCACCATGCCCG
9150 9160 9170 9180 9190 9200 9210

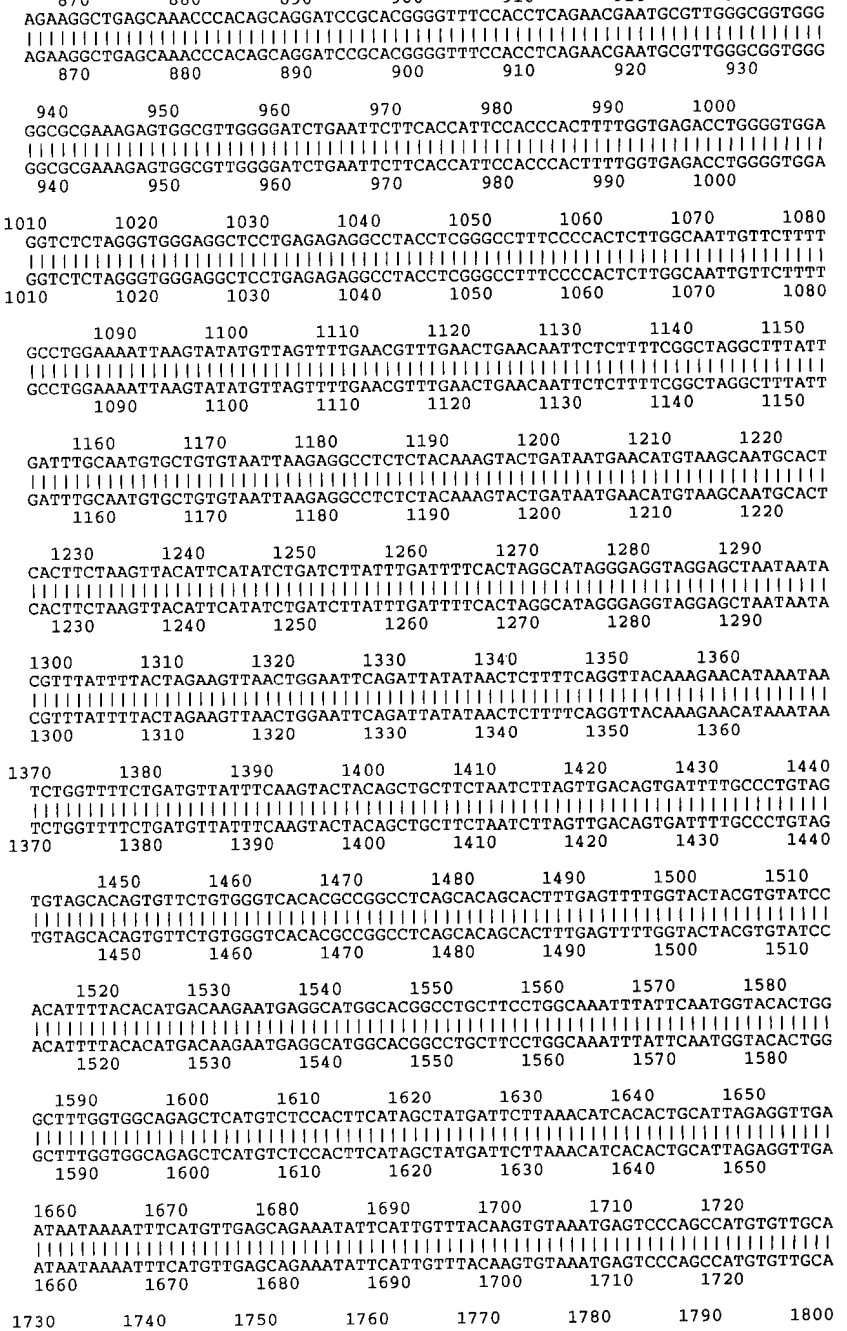


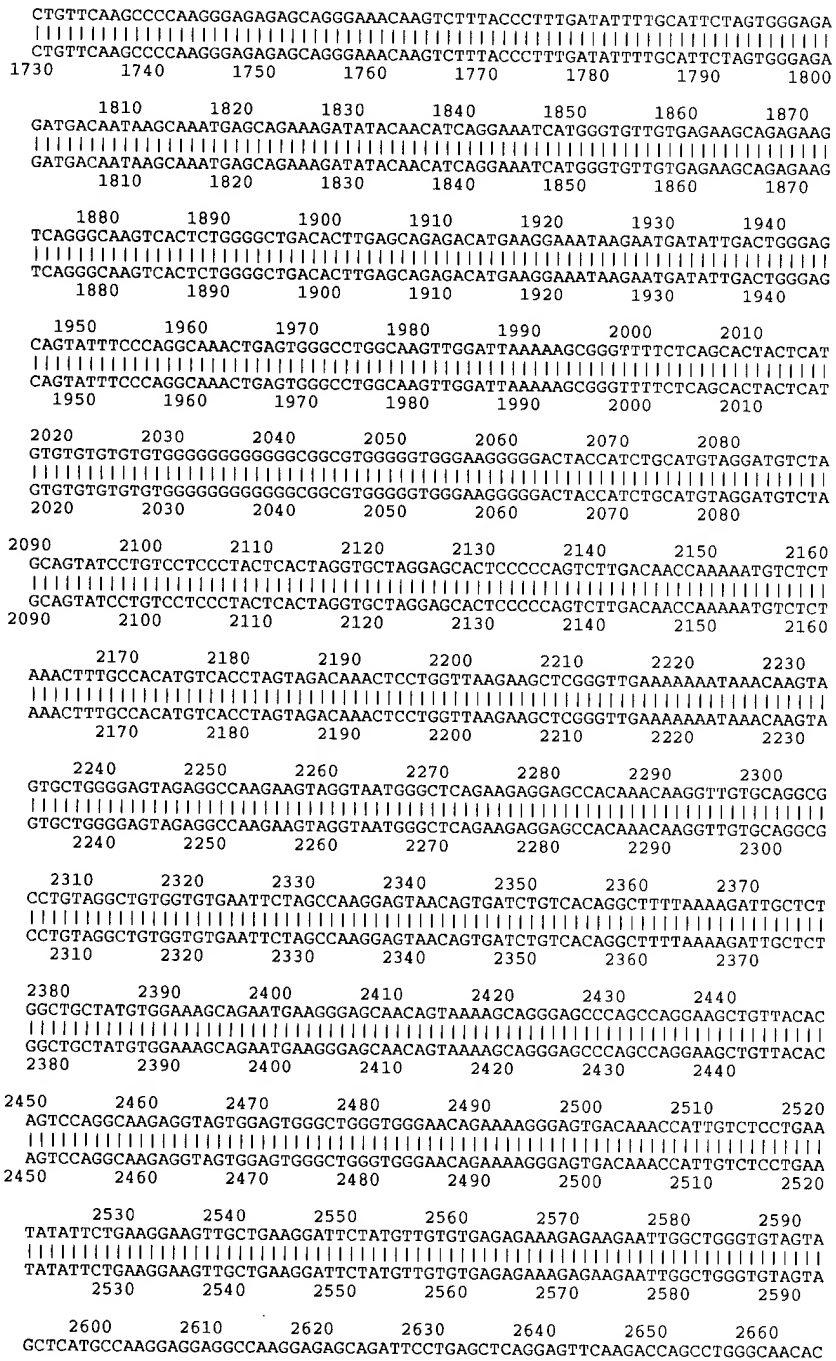


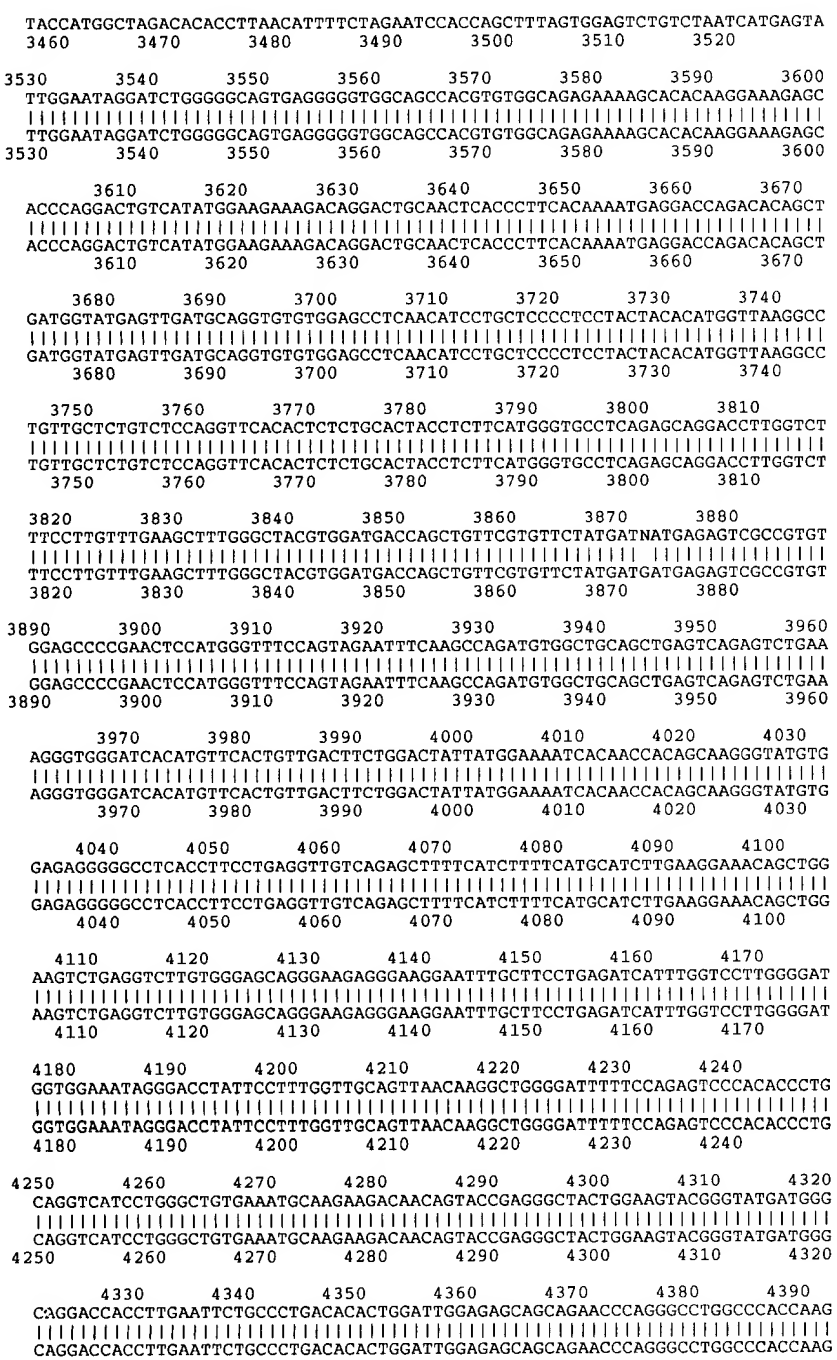
2. US-09-497-957-1 (1-10825)
US-09-497-957-5 Sequence 5, Application US/09497957

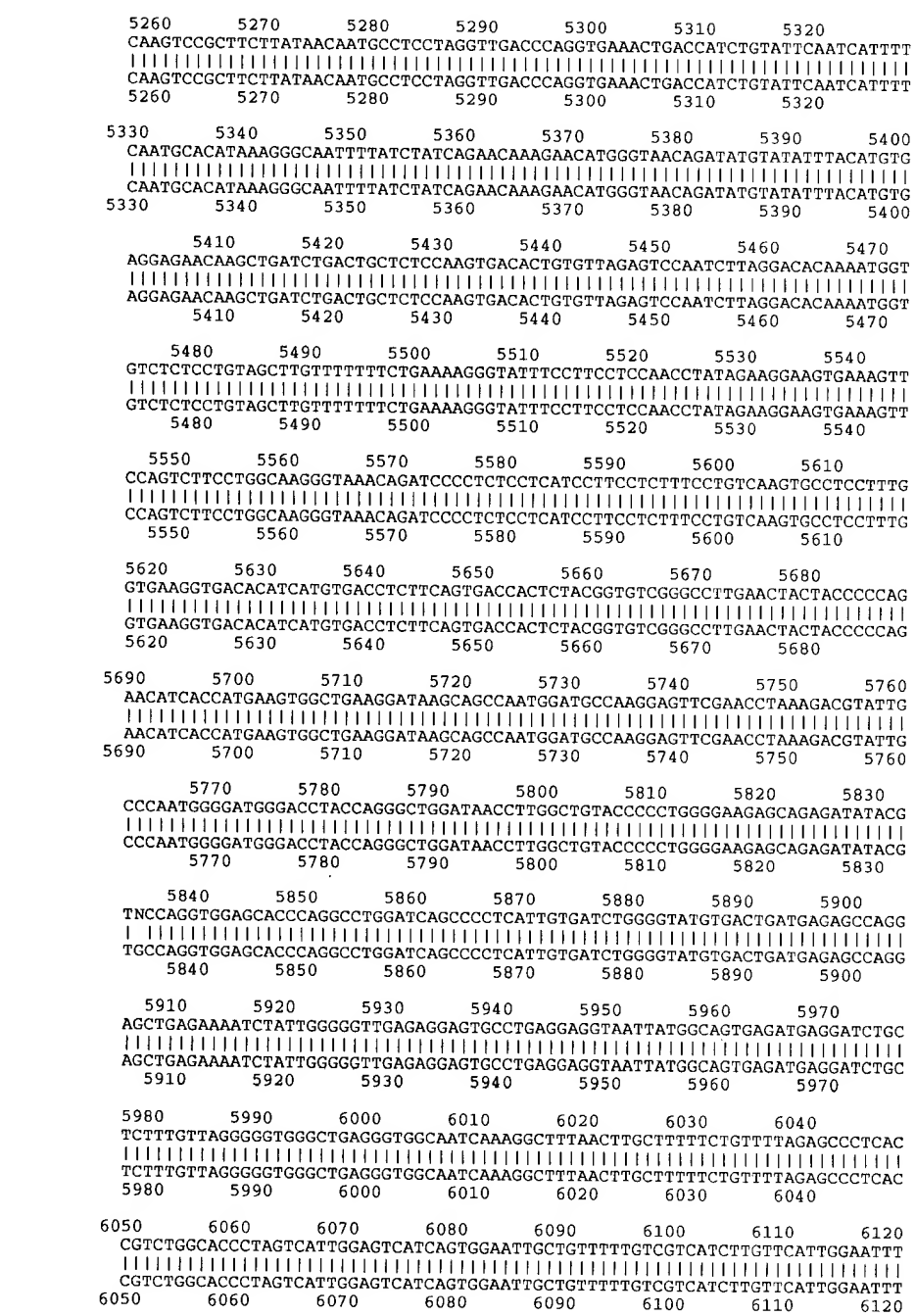
Initial Score	=	10825	Optimized Score	=	10825	Significance	=	4.02
Residue Identity	=	99%	Matches	=	10823	Mismatches	=	2
Gaps	=	0	Conservative Substitutions				=	0

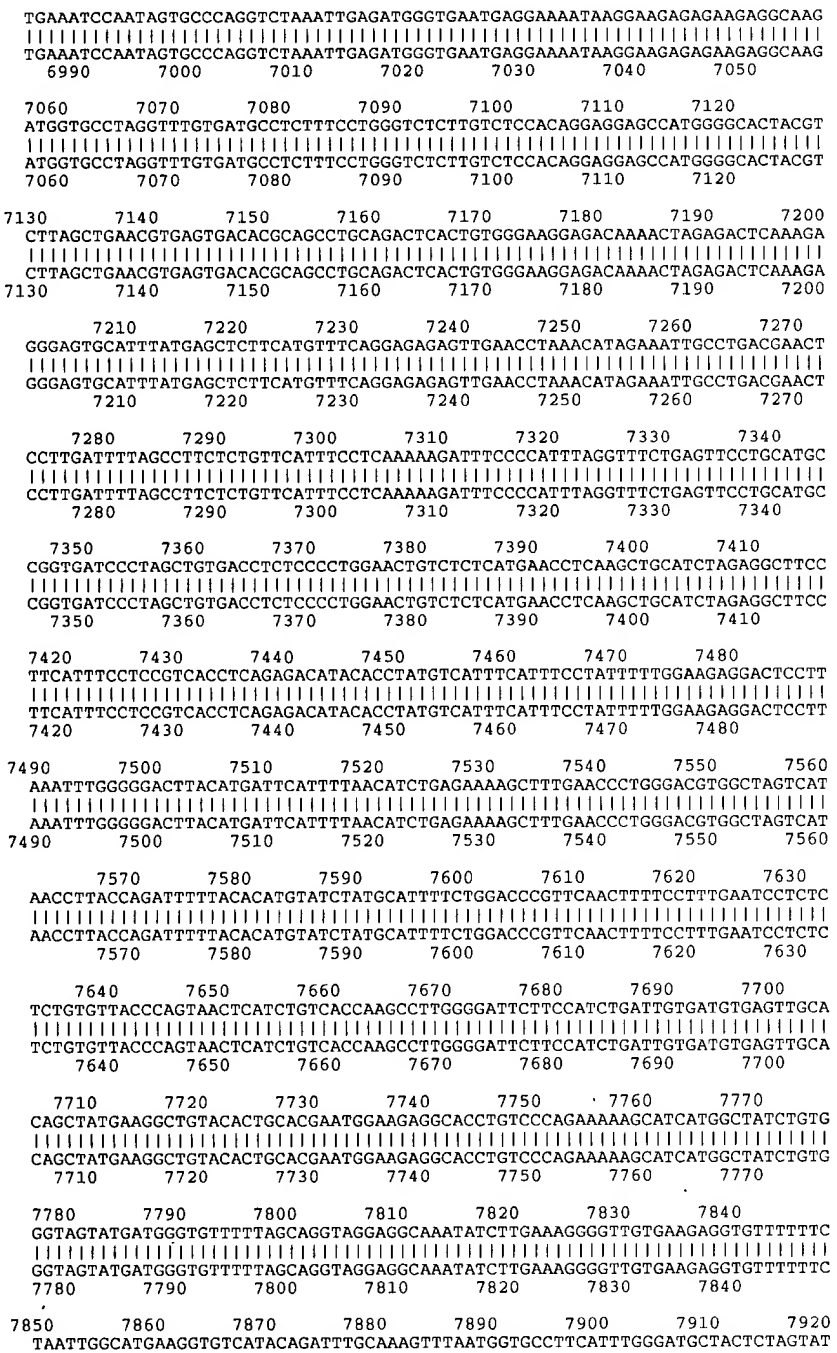


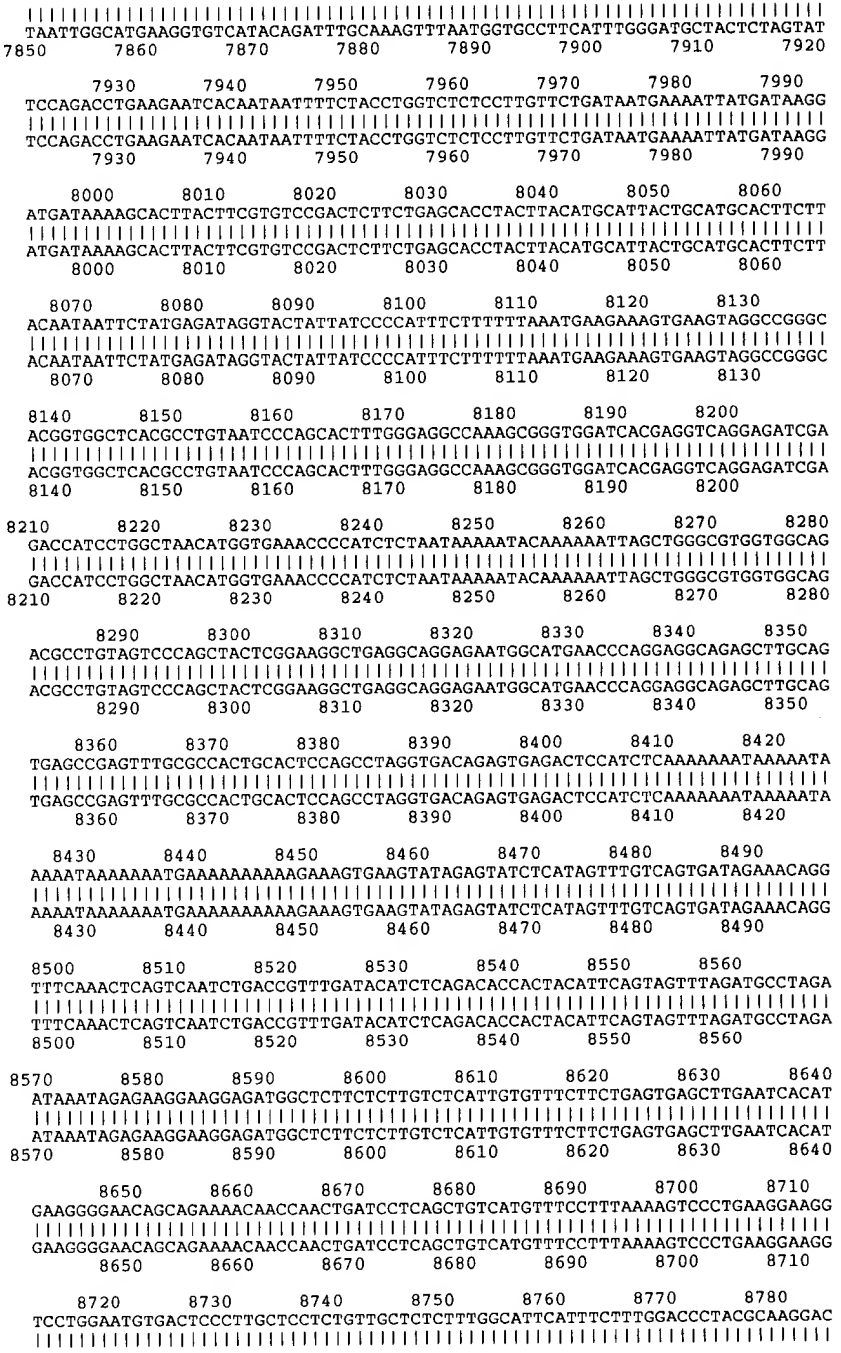


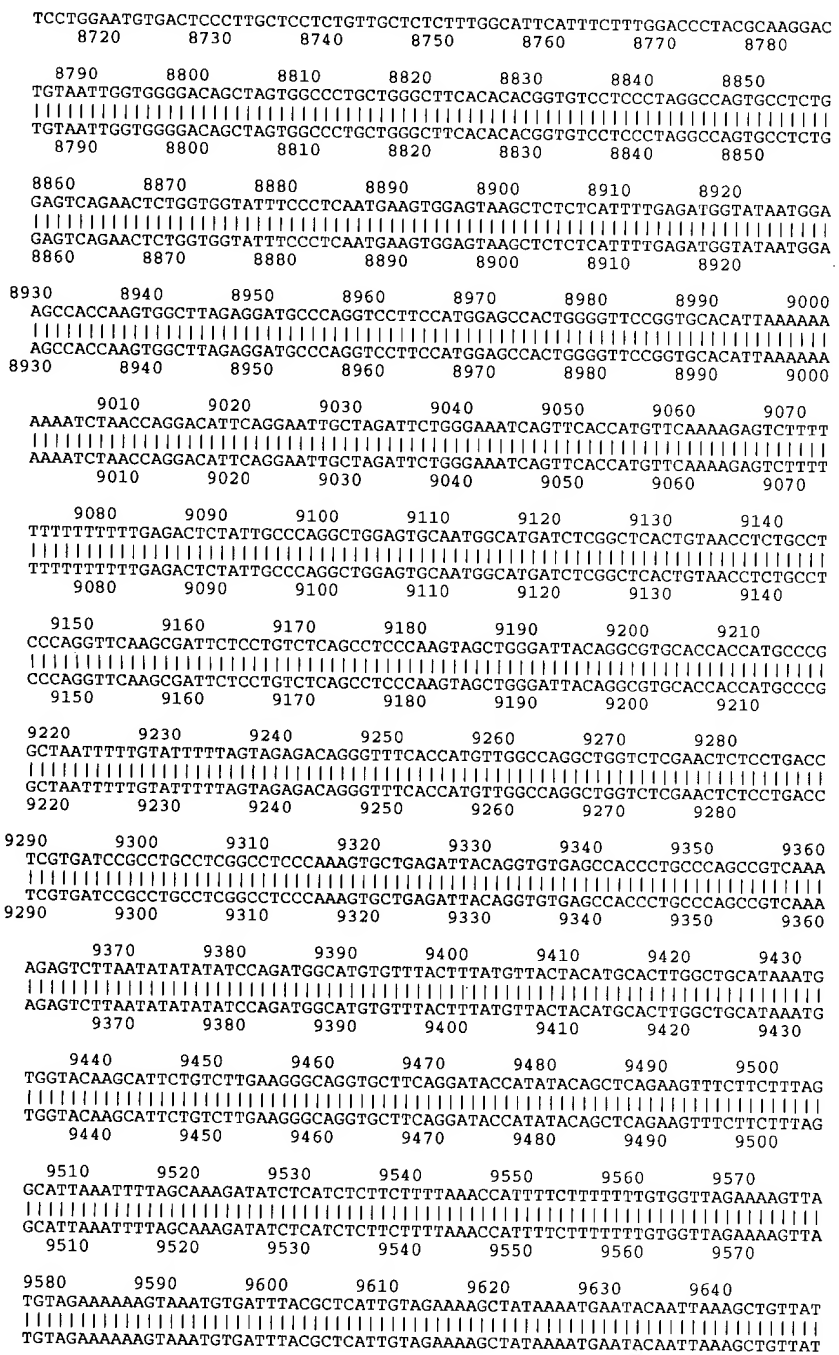

CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
1730 1740 1750 1760 1770 1780 1790 1800
1810 1820 1830 1840 1850 1860 1870
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
1810 1820 1830 1840 1850 1860 1870
1880 1890 1900 1910 1920 1930 1940
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
1880 1890 1900 1910 1920 1930 1940
1950 1960 1970 1980 1990 2000 2010
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
1950 1960 1970 1980 1990 2000 2010
2020 2030 2040 2050 2060 2070 2080
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
2020 2030 2040 2050 2060 2070 2080
2090 2100 2110 2120 2130 2140 2150 2160
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAATGTCTCT
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAATGTCTCT
2090 2100 2110 2120 2130 2140 2150 2160
2170 2180 2190 2200 2210 2220 2230
AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA
AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA
2170 2180 2190 2200 2210 2220 2230
2240 2250 2260 2270 2280 2290 2300
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
2240 2250 2260 2270 2280 2290 2300
2310 2320 2330 2340 2350 2360 2370
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
2310 2320 2330 2340 2350 2360 2370
2380 2390 2400 2410 2420 2430 2440
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
2380 2390 2400 2410 2420 2430 2440
2450 2460 2470 2480 2490 2500 2510 2520
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
2450 2460 2470 2480 2490 2500 2510 2520
2530 2540 2550 2560 2570 2580 2590
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
2530 2540 2550 2560 2570 2580 2590
2600 2610 2620 2630 2640 2650 2660
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
2600 2610 2620 2630 2640 2650 2660
2670 2680 2690 2700 2710 2720 2730
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
2670 2680 2690 2700 2710 2720 2730
2740 2750 2760 2770 2780 2790 2800
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
2740 2750 2760 2770 2780 2790 2800
2810 2820 2830 2840 2850 2860 2870 2880
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
2810 2820 2830 2840 2850 2860 2870 2880
2890 2900 2910 2920 2930 2940 2950
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
2890 2900 2910 2920 2930 2940 2950
2960 2970 2980 2990 3000 3010 3020
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
2960 2970 2980 2990 3000 3010 3020
3030 3040 3050 3060 3070 3080 3090
GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
3030 3040 3050 3060 3070 3080 3090
3100 3110 3120 3130 3140 3150 3160
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
3100 3110 3120 3130 3140 3150 3160
3170 3180 3190 3200 3210 3220 3230 3240
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
3170 3180 3190 3200 3210 3220 3230 3240
3250 3260 3270 3280 3290 3300 3310
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
3250 3260 3270 3280 3290 3300 3310
3320 3330 3340 3350 3360 3370 3380
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
3320 3330 3340 3350 3360 3370 3380
3390 3400 3410 3420 3430 3440 3450
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
3390 3400 3410 3420 3430 3440 3450
3460 3470 3480 3490 3500 3510 3520
TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA

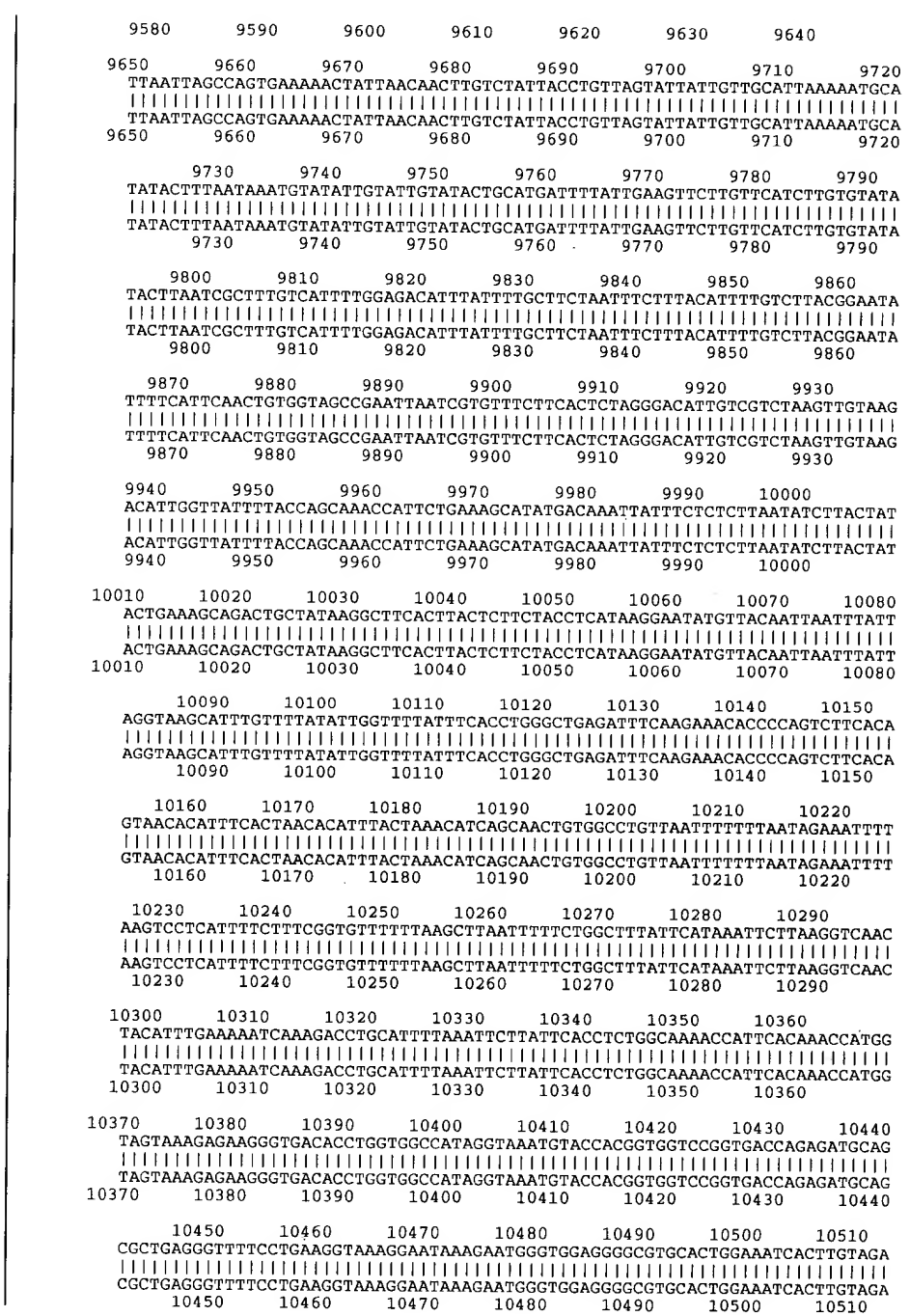

TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA
3460 3470 3480 3490 3500 3510 3520
3530 3540 3550 3560 3570 3580 3590 3600
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
3530 3540 3550 3560 3570 3580 3590 3600
3610 3620 3630 3640 3650 3660 3670
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
3610 3620 3630 3640 3650 3660 3670
3680 3690 3700 3710 3720 3730 3740
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
3680 3690 3700 3710 3720 3730 3740
3750 3760 3770 3780 3790 3800 3810
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
3750 3760 3770 3780 3790 3800 3810
3820 3830 3840 3850 3860 3870 3880
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGT
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGT
3820 3830 3840 3850 3860 3870 3880
3890 3900 3910 3920 3930 3940 3950 3960
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
3890 3900 3910 3920 3930 3940 3950 3960
3970 3980 3990 4000 4010 4020 4030
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
3970 3980 3990 4000 4010 4020 4030
4040 4050 4060 4070 4080 4090 4100
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
4040 4050 4060 4070 4080 4090 4100
4110 4120 4130 4140 4150 4160 4170
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
4110 4120 4130 4140 4150 4160 4170
4180 4190 4200 4210 4220 4230 4240
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
4180 4190 4200 4210 4220 4230 4240
4250 4260 4270 4280 4290 4300 4310 4320
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
4250 4260 4270 4280 4290 4300 4310 4320
4330 4340 4350 4360 4370 4380 4390
CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
4330 4340 4350 4360 4370 4380 4390
4400 4410 4420 4430 4440 4450 4460
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
4400 4410 4420 4430 4440 4450 4460
4470 4480 4490 4500 4510 4520 4530
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
4470 4480 4490 4500 4510 4520 4530
4540 4550 4560 4570 4580 4590 4600
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
4540 4550 4560 4570 4580 4590 4600
4610 4620 4630 4640 4650 4660 4670 4680
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
4610 4620 4630 4640 4650 4660 4670 4680
4690 4700 4710 4720 4730 4740 4750
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
4690 4700 4710 4720 4730 4740 4750
4760 4770 4780 4790 4800 4810 4820
CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGG
CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGG
4760 4770 4780 4790 4800 4810 4820
4830 4840 4850 4860 4870 4880 4890
CACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGTCACAAGGTCAGGAGTTTGA
CACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGTCACAAGGTCAGGAGTTTGA
4830 4840 4850 4860 4870 4880 4890
4900 4910 4920 4930 4940 4950 4960
GACCAGCCTGACCAACATGGTGAAACCCGTCTCTAAAAAAATACAAAAATTAGCTGGTCACAGTCATGCGCA
GACCAGCCTGACCAACATGGTGAAACCCGTCTCTAAAAAAATACAAAAATTAGCTGGTCACAGTCATGCGCA
4900 4910 4920 4930 4940 4950 4960
4970 4980 4990 5000 5010 5020 5030 5040
CCTGTAGTCCCAGCTAATTGGAAGGCTGAGGCAGGAGCATCGCTTGAACCTGGGAAGCGGAAGTTGCACTGA
CCTGTAGTCCCAGCTAATTGGAAGGCTGAGGCAGGAGCATCGCTTGAACCTGGGAAGCGGAAGTTGCACTGA
4970 4980 4990 5000 5010 5020 5030 5040
5050 5060 5070 5080 5090 5100 5110
GCCAAGATCGCGCCACTGCACTCCAGCCTAGGCAGCAGAGTGAGACTCCATCTTAAAAAAAAAAAAAAAAAA
GCCAAGATCGCGCCACTGCACTCCAGCCTAGGCAGCAGAGTGAGACTCCATCTTAAAAAAAAAAAAAAAAAA
5050 5060 5070 5080 5090 5100 5110
5120 5130 5140 5150 5160 5170 5180
AAAAAGAGAATTCAGAGATCTCAGCTATCATATGAATACCAGGACAAAATATCAAGTGAGGCCACTTATCAG
AAAAAGAGAATTCAGAGATCTCAGCTATCATATGAATACCAGGACAAAATATCAAGTGAGGCCACTTATCAG
5120 5130 5140 5150 5160 5170 5180
5190 5200 5210 5220 5230 5240 5250
AGTAGAAGAATCCTTTAGGTTAAAAGTTTCTTTCATAGAACATAGCAATAATCACTGAAGCTACCTATCTTA
AGTAGAAGAATCCTTTAGGTTAAAAGTTTCTTTCATAGAACATAGCAATAATCACTGAAGCTACCTATCTTA
5190 5200 5210 5220 5230 5240 5250

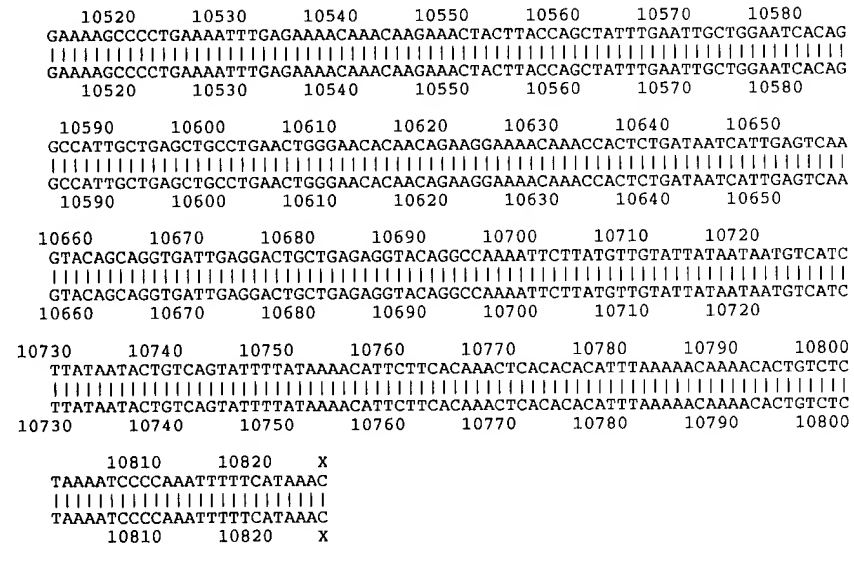

5260 5270 5280 5290 5300 5310 5320
CAAGTCCGCTTCTTATAACAATGCCTCCTAGGTTGACCCAGGTGAAACTGACCATCTGTATTCAATCATTTT
CAAGTCCGCTTCTTATAACAATGCCTCCTAGGTTGACCCAGGTGAAACTGACCATCTGTATTCAATCATTTT
5260 5270 5280 5290 5300 5310 5320
5330 5340 5350 5360 5370 5380 5390 5400
CAATGCACATAAAGGGCAATTTTATCTATCAGAACAAAGAACATGGGTAACAGATATGTATATTTACATGTG
CAATGCACATAAAGGGCAATTTTATCTATCAGAACAAAGAACATGGGTAACAGATATGTATATTTACATGTG
5330 5340 5350 5360 5370 5380 5390 5400
5410 5420 5430 5440 5450 5460 5470
AGGAGAACAAGCTGATCTGACTGCTCTCCAAGTGACACTGTGTTAGAGTCCAATCTTAGGACACAAAATGGT
AGGAGAACAAGCTGATCTGACTGCTCTCCAAGTGACACTGTGTTAGAGTCCAATCTTAGGACACAAAATGGT
5410 5420 5430 5440 5450 5460 5470
5480 5490 5500 5510 5520 5530 5540
GTCTCTCCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTT
GTCTCTCCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTT
5480 5490 5500 5510 5520 5530 5540
5550 5560 5570 5580 5590 5600 5610
CCAGTCTTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTG
CCAGTCTTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTG
5550 5560 5570 5580 5590 5600 5610
5620 5630 5640 5650 5660 5670 5680
GTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG
GTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG
5620 5630 5640 5650 5660 5670 5680
5690 5700 5710 5720 5730 5740 5750 5760
AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTG
AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTG
5690 5700 5710 5720 5730 5740 5750 5760
5770 5780 5790 5800 5810 5820 5830
CCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACG
CCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACG
5770 5780 5790 5800 5810 5820 5830
5840 5850 5860 5870 5880 5890 5900
TNCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGG
TGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGG
5840 5850 5860 5870 5880 5890 5900
5910 5920 5930 5940 5950 5960 5970
AGCTGAGAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGC
AGCTGAGAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGC
5910 5920 5930 5940 5950 5960 5970
5980 5990 6000 6010 6020 6030 6040
TCTTTGTTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCAC
TCTTTGTTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCAC
5980 5990 6000 6010 6020 6030 6040
6050 6060 6070 6080 6090 6100 6110 6120
CGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTT
CGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTT
6050 6060 6070 6080 6090 6100 6110 6120
6130 6140 6150 6160 6170 6180 6190
TGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGTGAGTAGGAACAAGGGGGAAGTCTCTTAGTACCTCTGC
TGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGTGAGTAGGAACAAGGGGGAAGTCTCTTAGTACCTCTGC
6130 6140 6150 6160 6170 6180 6190
6200 6210 6220 6230 6240 6250 6260
CCCAGGGCACAGTGGGAAGAGGGGCAGAGGGGATCTGGCATCCATGGGAAGCATTTTTCTCATTTATATTCT
CCCAGGGCACAGTGGGAAGAGGGGCAGAGGGGATCTGGCATCCATGGGAAGCATTTTTCTCATTTATATTCT
6200 6210 6220 6230 6240 6250 6260
6270 6280 6290 6300 6310 6320 6330
TTGGGGACACCAGCAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCCAGAATGAAAGTCTCTAATTCAAC
TTGGGGACACCAGCAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCCAGAATGAAAGTCTCTAATTCAAC
6270 6280 6290 6300 6310 6320 6330
6340 6350 6360 6370 6380 6390 6400
AAACATCTTCAGAGCACCTACTATTTTGCAAGAGCTGTTTAAGGTAGTACAGGGGCTTTGAGGTTGAGAAGT
AAACATCTTCAGAGCACCTACTATTTTGCAAGAGCTGTTTAAGGTAGTACAGGGGCTTTGAGGTTGAGAAGT
6340 6350 6360 6370 6380 6390 6400
6410 6420 6430 6440 6450 6460 6470 6480
CACTGTGGCTATTCTCAGAACCCAAATCTGGTAGGGAATGAAATTGATAGCAAGTAAATGTAGTTAAAGAAG
CACTGTGGCTATTCTCAGAACCCAAATCTGGTAGGGAATGAAATTGATAGCAAGTAAATGTAGTTAAAGAAG
6410 6420 6430 6440 6450 6460 6470 6480
6490 6500 6510 6520 6530 6540 6550
ACCCCATGAGGTCCTAAAGCAGGCAGGAAGCAAATGCTTAGGGTGTCAAAGGAAAGAATGATCACATTCAGC
ACCCCATGAGGTCCTAAAGCAGGCAGGAAGCAAATGCTTAGGGTGTCAAAGGAAAGAATGATCACATTCAGC
6490 6500 6510 6520 6530 6540 6550
6560 6570 6580 6590 6600 6610 6620
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
6560 6570 6580 6590 6600 6610 6620
6630 6640 6650 6660 6670 6680 6690
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
6630 6640 6650 6660 6670 6680 6690
6700 6710 6720 6730 6740 6750 6760
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
6700 6710 6720 6730 6740 6750 6760
6770 6780 6790 6800 6810 6820 6830 6840
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
6770 6780 6790 6800 6810 6820 6830 6840
6850 6860 6870 6880 6890 6900 6910
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
6850 6860 6870 6880 6890 6900 6910
6920 6930 6940 6950 6960 6970 6980
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT
6920 6930 6940 6950 6960 6970 6980
6990 7000 7010 7020 7030 7040 7050





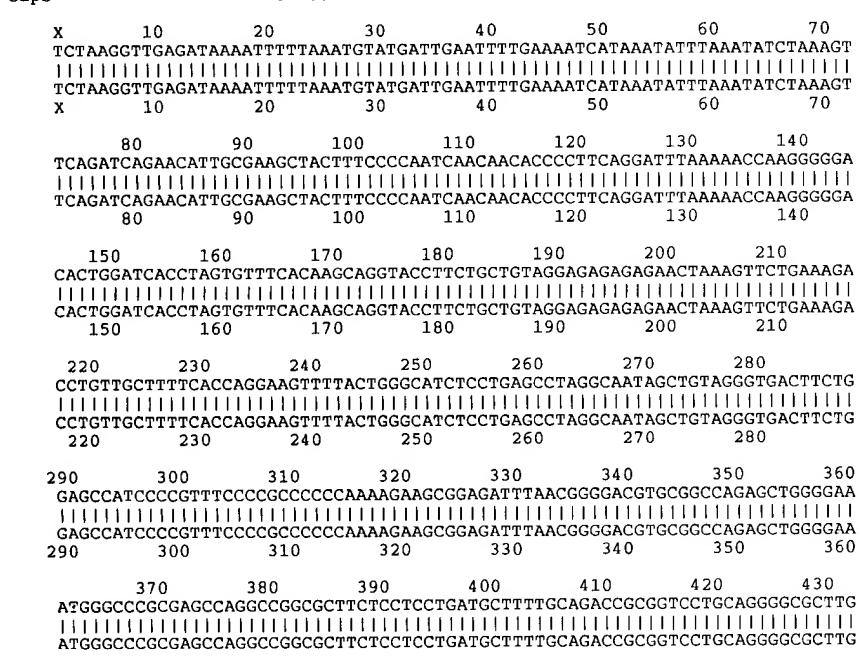

TCCTGGAATGTGACTCCCTTGCTCCTCTGTTGCTCTCTTTGGCATTCATTTCTTTGGACCCTACGCAAGGAC
8720 8730 8740 8750 8760 8770 8780
8790 8800 8810 8820 8830 8840 8850
TGTAATTGGTGGGGACAGCTAGTGGCCCTGCTGGGCTTCACACACGGTGTCCTCCCTAGGCCAGTGCCTCTG
8790 8800 8810 8820 8830 8840 8850
8860 8870 8880 8890 8900 8910 8920
GAGTCAGAACTCTGGTGGTATTTCCCTCAATGAAGTGGAGTAAGCTCTCTCATTTTGAGATGGTATAATGGA
8860 8870 8880 8890 8900 8910 8920
8930 8940 8950 8960 8970 8980 8990 9000
AGCCACCAAGTGGCTTAGAGGATGCCCAGGTCCTTCCATGGAGCCACTGGGGTTCCGGTGCACATTAAAAAA
8930 8940 8950 8960 8970 8980 8990 9000
9010 9020 9030 9040 9050 9060 9070
AAAATCTAACCAGGACATTCAGGAATTGCTAGATTCTGGGAAATCAGTTCACCATGTTCAAAAGAGTCTTTT
9010 9020 9030 9040 9050 9060 9070
9080 9090 9100 9110 9120 9130 9140
TTTTTTTTTTGAGACTCTATTGCCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGTAACCTCTGCCT
9080 9090 9100 9110 9120 9130 9140
9150 9160 9170 9180 9190 9200 9210
CCCAGGTTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCACCACCATGCCCG
9150 9160 9170 9180 9190 9200 9210
9220 9230 9240 9250 9260 9270 9280
GCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCTCCTGACC
9220 9230 9240 9250 9260 9270 9280
9290 9300 9310 9320 9330 9340 9350 9360
TCGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGAGATTACAGGTGTGAGCCACCCTGCCCAGCCGTCAAA
9290 9300 9310 9320 9330 9340 9350 9360
9370 9380 9390 9400 9410 9420 9430
AGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAATG
9370 9380 9390 9400 9410 9420 9430
9440 9450 9460 9470 9480 9490 9500
TGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCTTCTTTAG
9440 9450 9460 9470 9480 9490 9500
9510 9520 9530 9540 9550 9560 9570
GCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTTTTTTGTGGTTAGAAAAGTTA
9510 9520 9530 9540 9550 9560 9570
9580 9590 9600 9610 9620 9630 9640
TGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTAT

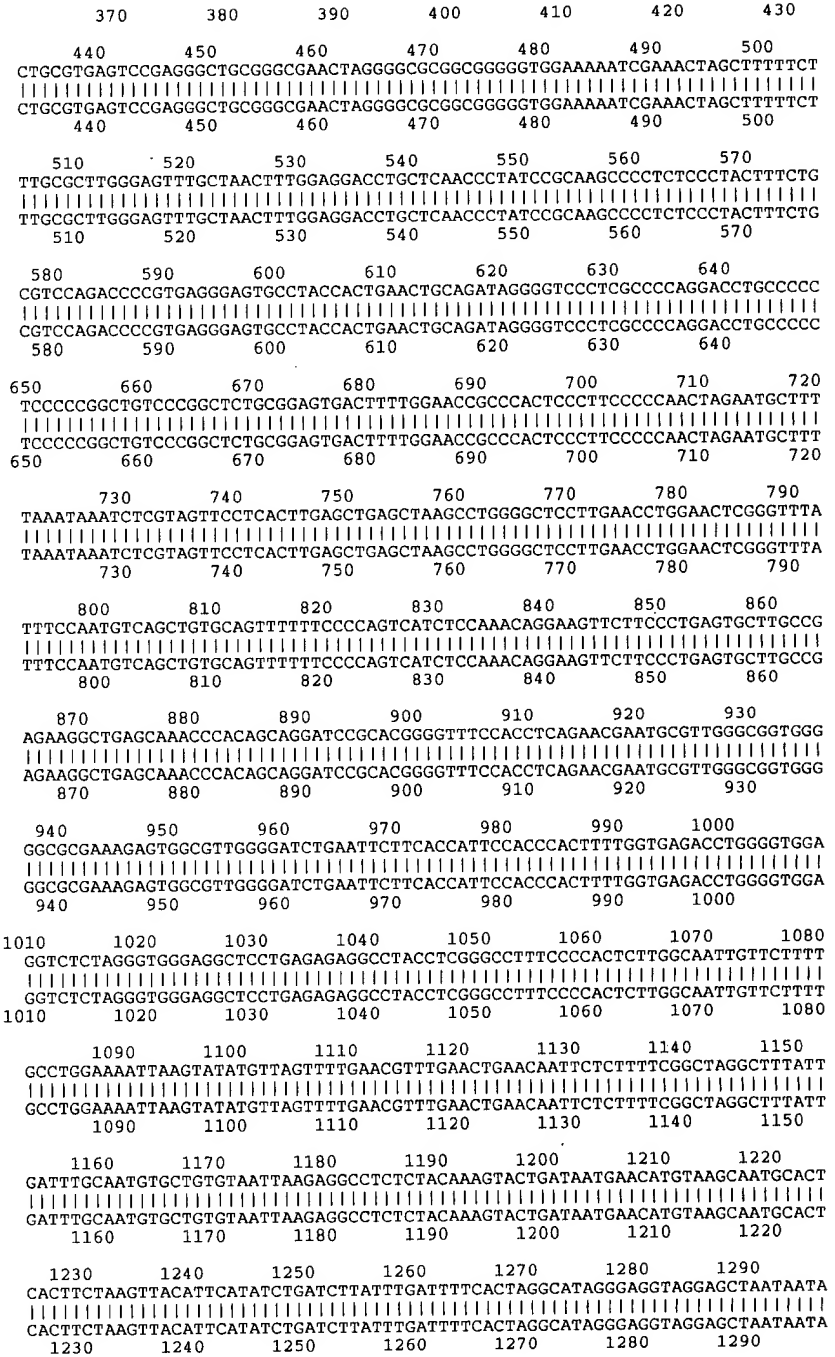

9580 9590 9600 9610 9620 9630 9640
9650 9660 9670 9680 9690 9700 9710 9720
TTAATTAGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCA
9650 9660 9670 9680 9690 9700 9710 9720
9730 9740 9750 9760 9770 9780 9790
TATACTTTAATAAATGTATATTGTATTGTATACTGCATGATTTTATTGAAGTTCTTGTTCATCTTGTGTATA
9730 9740 9750 9760 9770 9780 9790
9800 9810 9820 9830 9840 9850 9860
TACTTAATCGCTTTGTCATTTTGGAGACATTTATTTTGCTTCTAATTTCTTTACATTTTGTCTTACGGAATA
9800 9810 9820 9830 9840 9850 9860
9870 9880 9890 9900 9910 9920 9930
TTTTCATTCAACTGTGGTAGCCGAATTAATCGTGTTTCTTCACTCTAGGGACATTGTCGTCTAAGTTGTAAG
9870 9880 9890 9900 9910 9920 9930
9940 9950 9960 9970 9980 9990 10000
ACATTGGTTATTTTACCAGCAAACCATTCTGAAAGCATATGACAAATTATTTCTCTCTTAATATCTTACTAT
9940 9950 9960 9970 9980 9990 10000
10010 10020 10030 10040 10050 10060 10070 10080
ACTGAAAGCAGACTGCTATAAGGCTTCACTTACTCTTCTACCTCATAAGGAATATGTTACAATTAATTTATT
10010 10020 10030 10040 10050 10060 10070 10080
10090 10100 10110 10120 10130 10140 10150
AGGTAAGCATTTGTTTTATATTGGTTTTATTTCACCTGGGCTGAGATTTCAAGAAACACCCCAGTCTTCACA
10090 10100 10110 10120 10130 10140 10150
10160 10170 10180 10190 10200 10210 10220
GTAACACATTTCACTAACACATTTACTAAACATCAGCAACTGTGGCCTGTTAATTTTTTTAATAGAAATTTT
10160 10170 10180 10190 10200 10210 10220
10230 10240 10250 10260 10270 10280 10290
AAGTCCTCATTTTCTTTCGGTGTTTTTTAAGCTTAATTTTTCTGGCTTTATTCATAAATTCTTAAGGTCAAC
10230 10240 10250 10260 10270 10280 10290
10300 10310 10320 10330 10340 10350 10360
TACATTTGAAAAATCAAAGACCTGCATTTTAAATTCTTATTCACCTCTGGCAAAACCATTCACAAACCATGG
10300 10310 10320 10330 10340 10350 10360
10370 10380 10390 10400 10410 10420 10430 10440
TAGTAAAGAGAAGGGTGACACCTGGTGGCCATAGGTAAATGTACCACGGTGGTCCGGTGACCAGAGATGCAG
10370 10380 10390 10400 10410 10420 10430 10440
10450 10460 10470 10480 10490 10500 10510
CGCTGAGGGTTTTCCTGAAGGTAAAGGAATAAACAATGGGTGGAGGGGCGTGCACTGGAAATCACTTGTAGA
10450 10460 10470 10480 10490 10500 10510

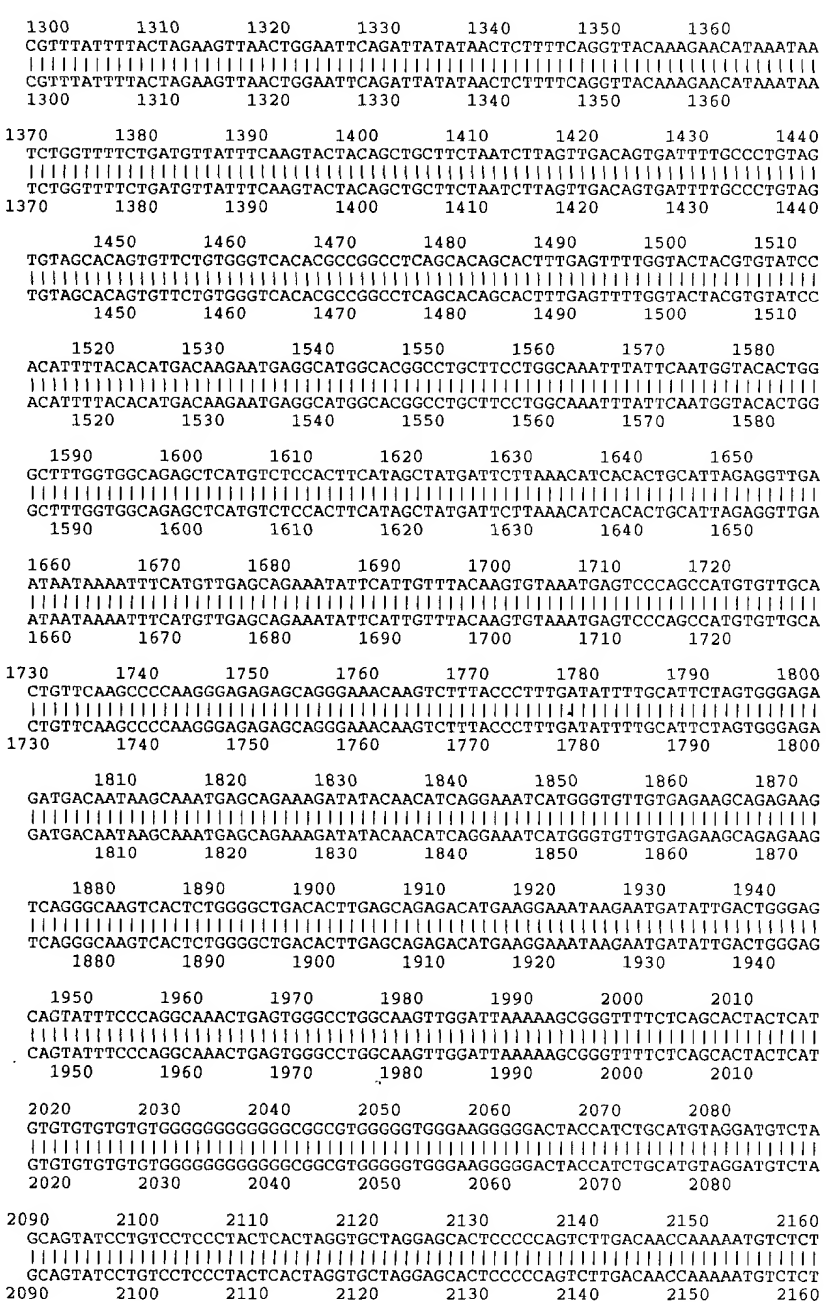


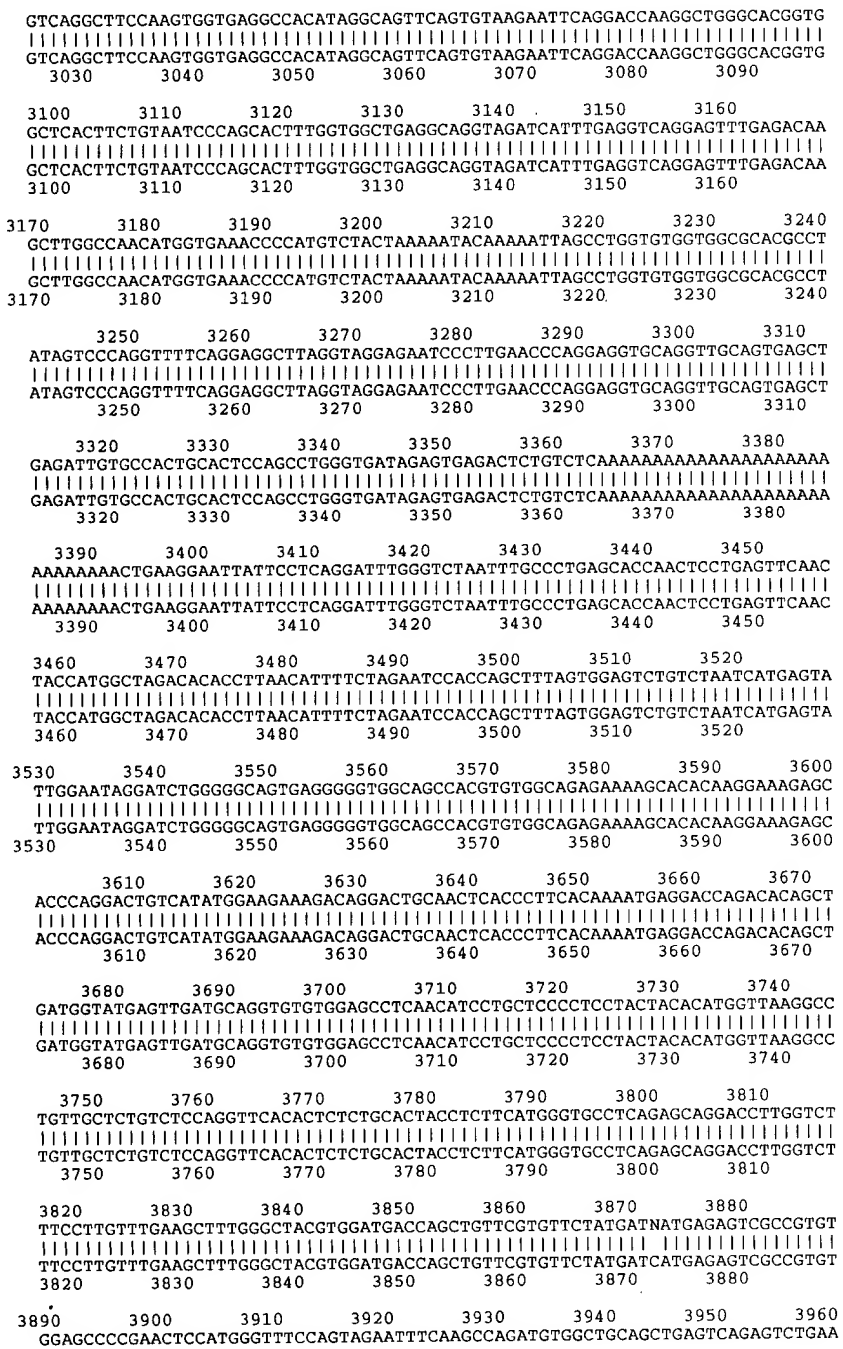
3. US-09-497-957-1 (1-10825)
US-09-497-957-3 Sequence 3, Application US/09497957

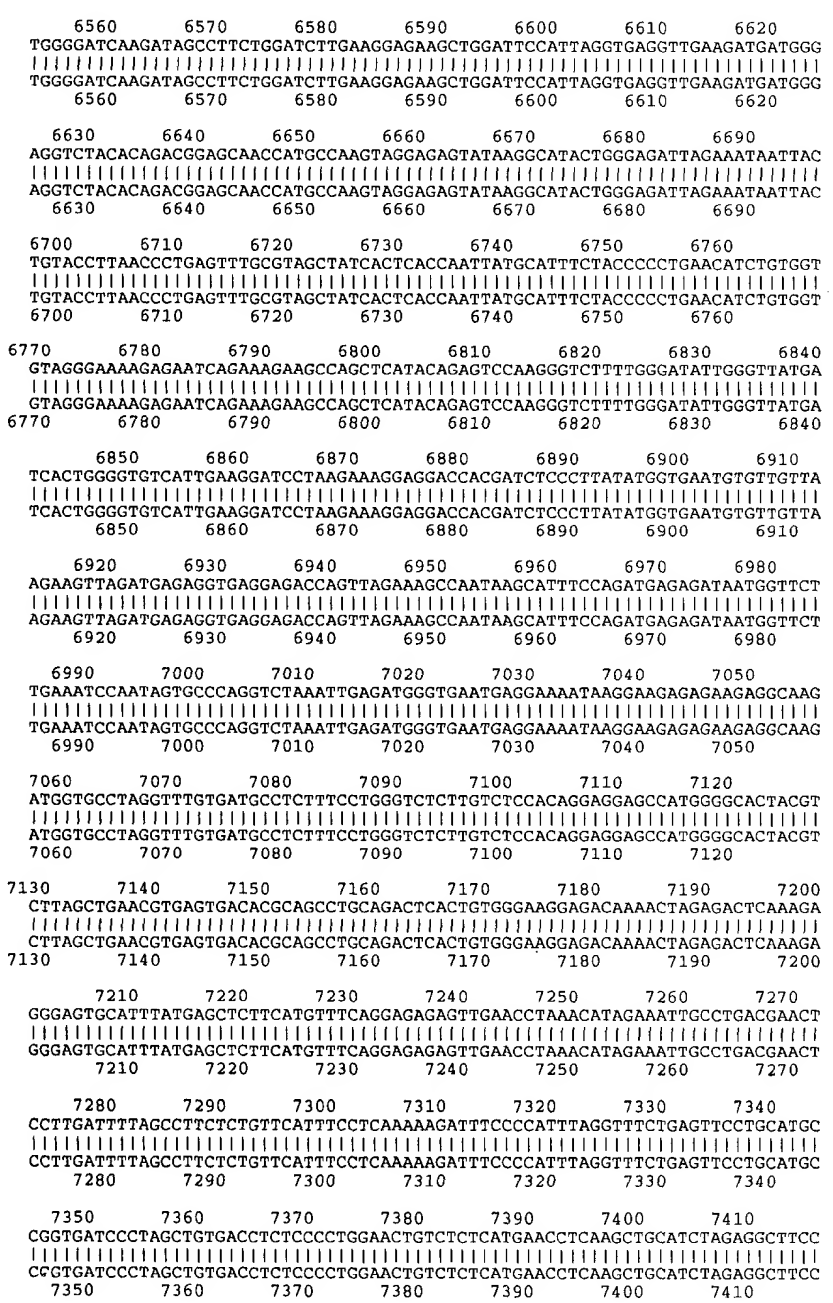
Initial Score	=	10825	Optimized Score	=	10825	Significance =	4.02
Residue Identity	=	99%	Matches	=	10823	Mismatches =	2
Gaps	=	0	Conservative Substitutions			=	0

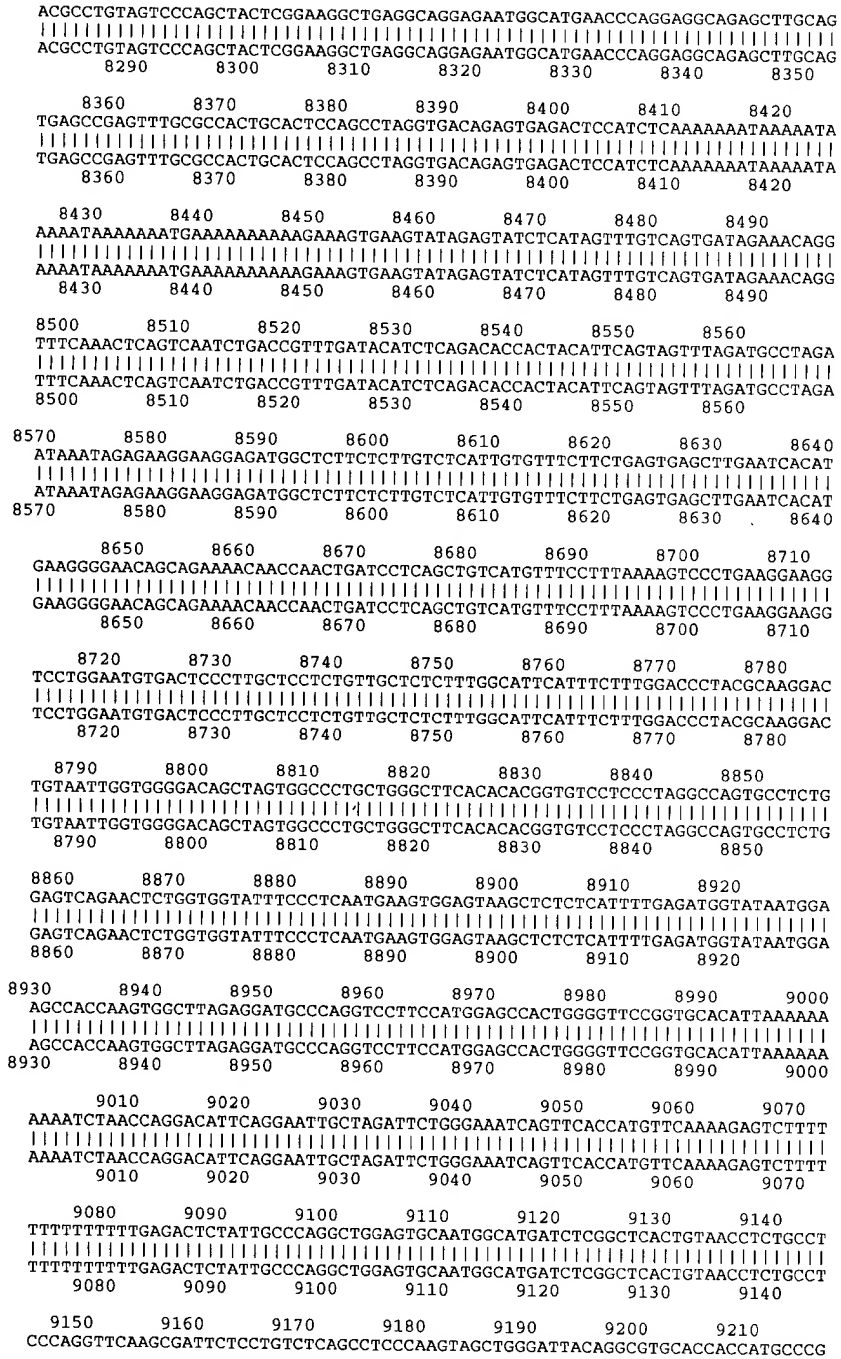


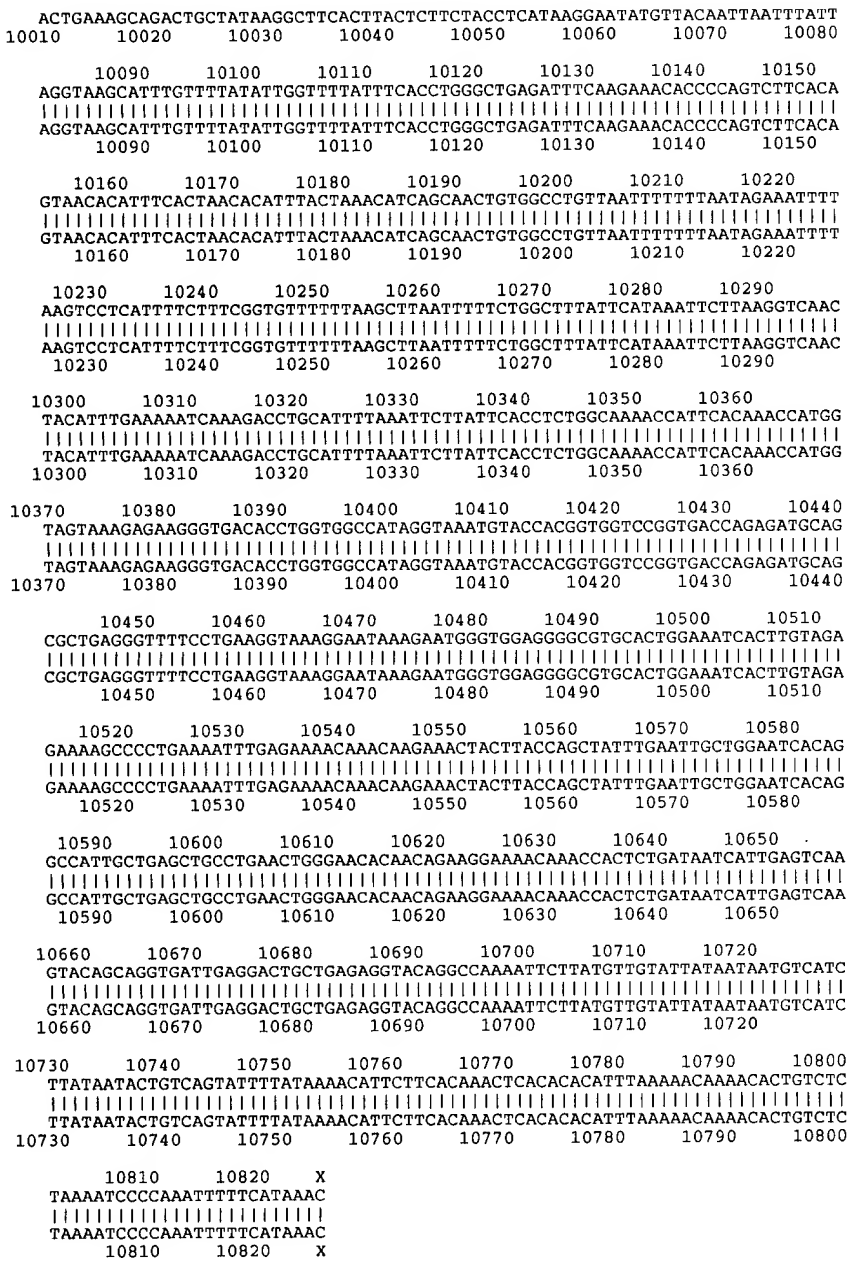



1300 1310 1320 1330 1340 1350 1360
CGTTTATTTTACTAGAAGTTAACTGGAATTCAGATTATATAACTCTTTTCAGGTTACAAAGAACATAAATAA
CGTTTATTTTACTAGAAGTTAACTGGAATTCAGATTATATAACTCTTTTCAGGTTACAAAGAACATAAATAA
1300 1310 1320 1330 1340 1350 1360
1370 1380 1390 1400 1410 1420 1430 1440
TCTGGTTTTCTGATGTTATTTCAAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTGATTTTGCCCCTGTAG
TCTGGTTTTCTGATGTTATTTCAAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTGATTTTGCCCTGTAG
1370 1380 1390 1400 1410 1420 1430 1440
1450 1460 1470 1480 1490 1500 1510
TGTAGCACAGTGTTCTGTGGGTCACACGCCGGCCTCAGCACAGCACTTTGAGTTTTGGTACTACGTGTATCC
TGTAGCACAGTGTTCTGTGGGTCACACGCCGGCCTCAGCACAGCACTTTGAGTTTTGGTACTACGTGTATCC
1450 1460 1470 1480 1490 1500 1510
1520 1530 1540 1550 1560 1570 1580
ACATTTTACACATGACAAGAATGAGGCATGGCACGGCCTGCTTCCTGGCAAATTTATTCAATGGTACACTGG
ACATTTTACACATGACAAGAATGAGGCATGGCACGGCCTGCTTCCTGGCAAATTTATTCAATGGTACACTGG
1520 1530 1540 1550 1560 1570 1580
1590 1600 1610 1620 1630 1640 1650
GCTTTGGTGGCAGAGCTCATGTCTCCACTTCATAGCTATGATTCTTAAACATCACACTGCATTAGAGGTTGA
GCTTTGGTGGCAGAGCTCATGTCTCCACTTCATAGCTATGATTCTTAAACATCACACTGCATTAGAGGTTGA
1590 1600 1610 1620 1630 1640 1650
1660 1670 1680 1690 1700 1710 1720
ATAATAAAATTTCATGTTGAGCAGAAATATTCATTGTTTACAAGTGTAAATGAGTCCCAGCCATGTGTTGCA
ATAATAAAATTTCATGTTGAGCAGAAATATTCATTGTTTACAAGTGTAAATGAGTCCCAGCCATGTGTTGCA
1660 1670 1680 1690 1700 1710 1720
1730 1740 1750 1760 1770 1780 1790 1800
CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
1730 1740 1750 1760 1770 1780 1790 1800
1810 1820 1830 1840 1850 1860 1870
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
1810 1820 1830 1840 1850 1860 1870
1880 1890 1900 1910 1920 1930 1940
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
1880 1890 1900 1910 1920 1930 1940
1950 1960 1970 1980 1990 2000 2010
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
1950 1960 1970 1980 1990 2000 2010
2020 2030 2040 2050 2060 2070 2080
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
2020 2030 2040 2050 2060 2070 2080
2090 2100 2110 2120 2130 2140 2150 2160
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAATGTCTCT
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAATGTCTCT
2090 2100 2110 2120 2130 2140 2150 2160
2170 2180 2190 2200 2210 2220 2230
AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA
AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA
2170 2180 2190 2200 2210 2220 2230
2240 2250 2260 2270 2280 2290 2300
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
2240 2250 2260 2270 2280 2290 2300
2310 2320 2330 2340 2350 2360 2370
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
2310 2320 2330 2340 2350 2360 2370
2380 2390 2400 2410 2420 2430 2440
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
2380 2390 2400 2410 2420 2430 2440
2450 2460 2470 2480 2490 2500 2510 2520
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
2450 2460 2470 2480 2490 2500 2510 2520
2530 2540 2550 2560 2570 2580 2590
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
2530 2540 2550 2560 2570 2580 2590
2600 2610 2620 2630 2640 2650 2660
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
2600 2610 2620 2630 2640 2650 2660
2670 2680 2690 2700 2710 2720 2730
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
2670 2680 2690 2700 2710 2720 2730
2740 2750 2760 2770 2780 2790 2800
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
2740 2750 2760 2770 2780 2790 2800
2810 2820 2830 2840 2850 2860 2870 2880
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
2810 2820 2830 2840 2850 2860 2870 2880
2890 2900 2910 2920 2930 2940 2950
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
2890 2900 2910 2920 2930 2940 2950
2960 2970 2980 2990 3000 3010 3020
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
2960 2970 2980 2990 3000 3010 3020
3030 3040 3050 3060 3070 3080 3090

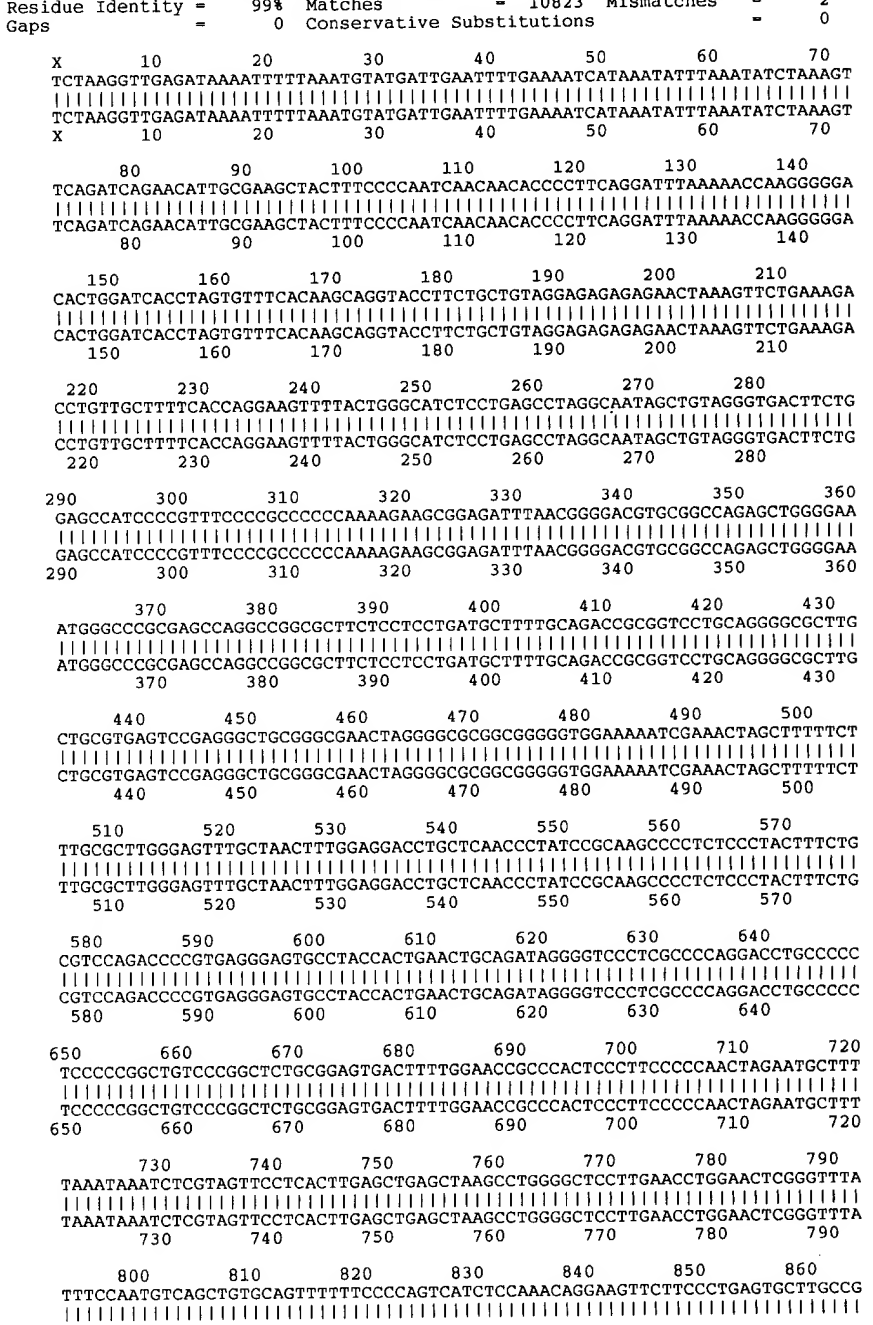

GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
3030 3040 3050 3060 3070 3080 3090
3100 3110 3120 3130 3140 3150 3160
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
3100 3110 3120 3130 3140 3150 3160
3170 3180 3190 3200 3210 3220 3230 3240
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
3170 3180 3190 3200 3210 3220 3230 3240
3250 3260 3270 3280 3290 3300 3310
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
3250 3260 3270 3280 3290 3300 3310
3320 3330 3340 3350 3360 3370 3380
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
3320 3330 3340 3350 3360 3370 3380
3390 3400 3410 3420 3430 3440 3450
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
3390 3400 3410 3420 3430 3440 3450
3460 3470 3480 3490 3500 3510 3520
TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA
TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA
3460 3470 3480 3490 3500 3510 3520
3530 3540 3550 3560 3570 3580 3590 3600
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
3530 3540 3550 3560 3570 3580 3590 3600
3610 3620 3630 3640 3650 3660 3670
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
3610 3620 3630 3640 3650 3660 3670
3680 3690 3700 3710 3720 3730 3740
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
3680 3690 3700 3710 3720 3730 3740
3750 3760 3770 3780 3790 3800 3810
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
3750 3760 3770 3780 3790 3800 3810
3820 3830 3840 3850 3860 3870 3880
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATNATGAGAGTCGCCGTGT
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGT
3820 3830 3840 3850 3860 3870 3880
3890 3900 3910 3920 3930 3940 3950 3960
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
3890 3900 3910 3920 3930 3940 3950 3960
3970 3980 3990 4000 4010 4020 4030
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
3970 3980 3990 4000 4010 4020 4030
4040 4050 4060 4070 4080 4090 4100
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
4040 4050 4060 4070 4080 4090 4100
4110 4120 4130 4140 4150 4160 4170
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
4110 4120 4130 4140 4150 4160 4170
4180 4190 4200 4210 4220 4230 4240
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
4180 4190 4200 4210 4220 4230 4240
4250 4260 4270 4280 4290 4300 4310 4320
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
4250 4260 4270 4280 4290 4300 4310 4320
4330 4340 4350 4360 4370 4380 4390
CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
4330 4340 4350 4360 4370 4380 4390
4400 4410 4420 4430 4440 4450 4460
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
4400 4410 4420 4430 4440 4450 4460
4470 4480 4490 4500 4510 4520 4530
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
4470 4480 4490 4500 4510 4520 4530
4540 4550 4560 4570 4580 4590 4600
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
4540 4550 4560 4570 4580 4590 4600
4610 4620 4630 4640 4650 4660 4670 4680
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
4610 4620 4630 4640 4650 4660 4670 4680
4690 4700 4710 4720 4730 4740 4750
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
4690 4700 4710 4720 4730 4740 4750
4760 4770 4780 4790 4800 4810 4820
CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGG

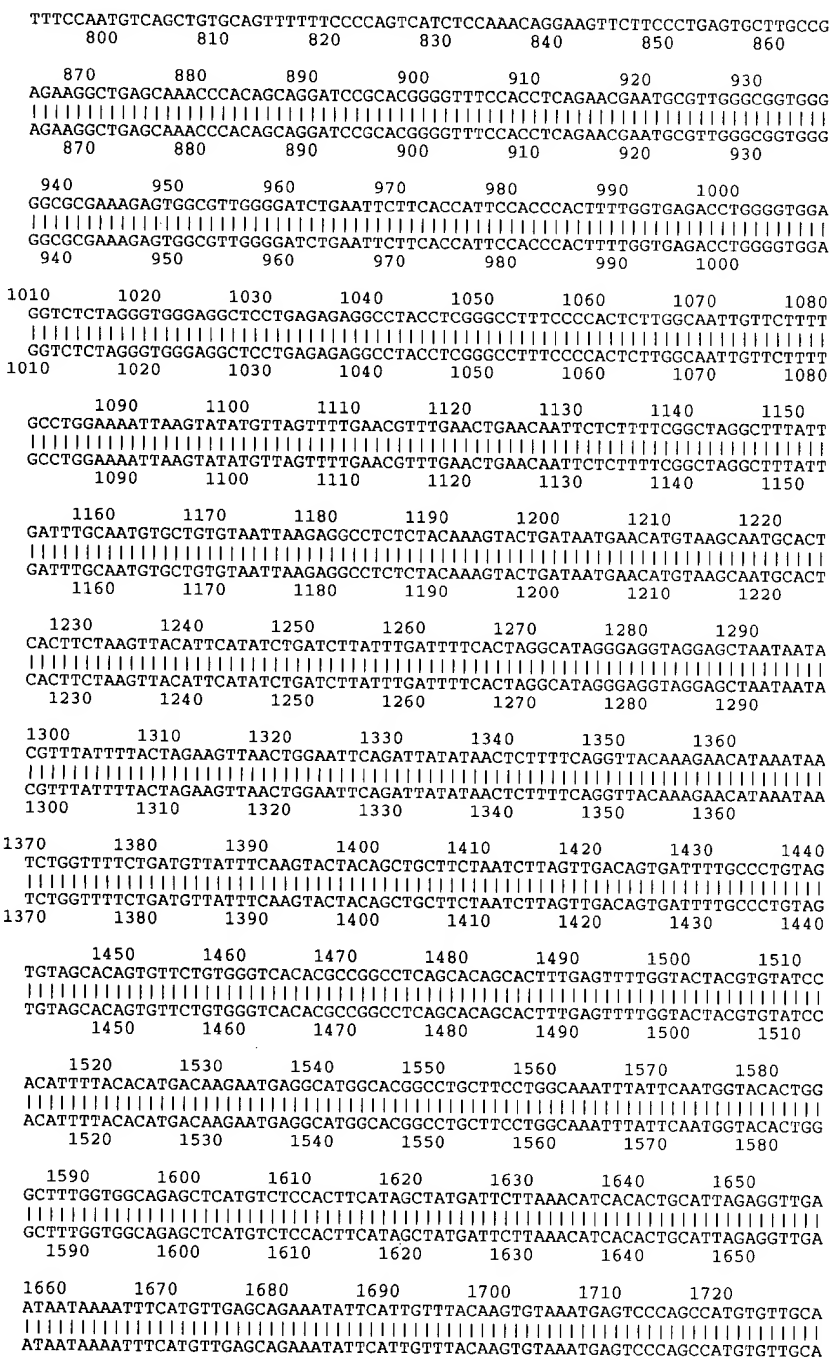

6560 6570 6580 6590 6600 6610 6620
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
6560 6570 6580 6590 6600 6610 6620
6630 6640 6650 6660 6670 6680 6690
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
6630 6640 6650 6660 6670 6680 6690
6700 6710 6720 6730 6740 6750 6760
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
6700 6710 6720 6730 6740 6750 6760
6770 6780 6790 6800 6810 6820 6830 6840
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
6770 6780 6790 6800 6810 6820 6830 6840
6850 6860 6870 6880 6890 6900 6910
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
6850 6860 6870 6880 6890 6900 6910
6920 6930 6940 6950 6960 6970 6980
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT
6920 6930 6940 6950 6960 6970 6980
6990 7000 7010 7020 7030 7040 7050
TGAAATCCAATAGTGCCCAGGTCTAAATTGAGATGGGTGAATGAGGAAAATAAGGAAGAGAGAAGAGGCAAG
TGAAATCCAATAGTGCCCAGGTCTAAATTGAGATGGGTGAATGAGGAAAATAAGGAAGAGAGAAGAGGCAAG
6990 7000 7010 7020 7030 7040 7050
7060 7070 7080 7090 7100 7110 7120
ATGGTGCCTAGGTTTGTGATGCCTCTTTCCTGGGTCTCTTGTCTCCACAGGAGGAGCCATGGGGCACTACGT
ATGGTGCCTAGGTTTGTGATGCCTCTTTCCTGGGTCTCTTGTCTCCACAGGAGGAGCCATGGGGCACTACGT
7060 7070 7080 7090 7100 7110 7120
7130 7140 7150 7160 7170 7180 7190 7200
CTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA
CTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA
7130 7140 7150 7160 7170 7180 7190 7200
7210 7220 7230 7240 7250 7260 7270
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACT
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACT
7210 7220 7230 7240 7250 7260 7270
7280 7290 7300 7310 7320 7330 7340
CCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGC
CCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGC
7280 7290 7300 7310 7320 7330 7340
7350 7360 7370 7380 7390 7400 7410
CGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCC
CGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCC
7350 7360 7370 7380 7390 7400 7410
7420 7430 7440 7450 7460 7470 7480
TTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTT
TTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTT
7420 7430 7440 7450 7460 7470 7480
7490 7500 7510 7520 7530 7540 7550 7560
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCAT
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCAT
7490 7500 7510 7520 7530 7540 7550 7560
7570 7580 7590 7600 7610 7620 7630
AACCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTC
AACCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTC
7570 7580 7590 7600 7610 7620 7630
7640 7650 7660 7670 7680 7690 7700
TCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCA
TCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCA
7640 7650 7660 7670 7680 7690 7700
7710 7720 7730 7740 7750 7760 7770
CAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTG
CAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTG
7710 7720 7730 7740 7750 7760 7770
7780 7790 7800 7810 7820 7830 7840
GGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGTGTTTTTTC
GGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGTGTTTTTTC
7780 7790 7800 7810 7820 7830 7840
7850 7860 7870 7880 7890 7900 7910 7920
TAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTAT
TAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTAT
7850 7860 7870 7880 7890 7900 7910 7920
7930 7940 7950 7960 7970 7980 7990
TCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGG
TCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGG
7930 7940 7950 7960 7970 7980 7990
8000 8010 8020 8030 8040 8050 8060
ATGATAAAAGCACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTT
ATGATAAAAGCACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTT
8000 8010 8020 8030 8040 8050 8060
8070 8080 8090 8100 8110 8120 8130
ACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGC
ACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGC
8070 8080 8090 8100 8110 8120 8130
8140 8150 8160 8170 8180 8190 8200
ACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGCGGGTGGATCACGAGGTCAGGAGATCGA
ACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAAGCGGGTGGATCACGAGGTCAGGAGATCGA
8140 8150 8160 8170 8180 8190 8200
8210 8220 8230 8240 8250 8260 8270 8280
GACCATCCTGGCTAACATGGTGAAACCCCATCTCTAATAAAAATACAAAAAATTAGCTGGGCGTGGTGGCAG
GACCATCCTGGCTAACATGGTGAAACCCCATCTCTAATAAAAATACAAAAAATTAGCTGGGCGTGGTGGCAG
8210 8220 8230 8240 8250 8260 8270 8280
8290 8300 8310 8320 8330 8340 8350

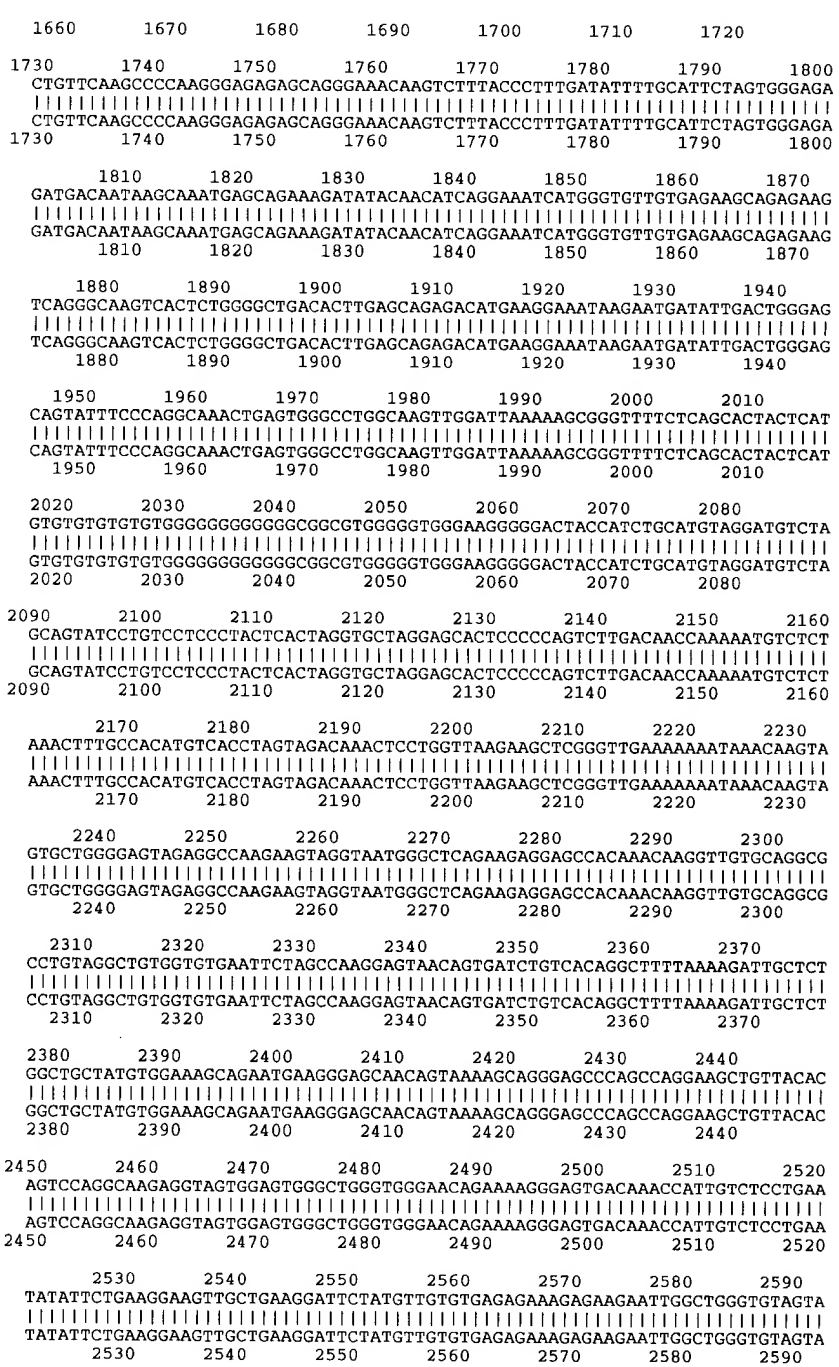

ACGCCTGTAGTCCCAGCTACTCGGAAGGCTGAGGCAGGAGAATGGCATGAACCCAGGAGGCAGAGCTTGCAG
ACGCCTGTAGTCCCAGCTACTCGGAAGGCTGAGGCAGGAGAATGGCATGAACCCAGGAGGCAGAGCTTGCAG
8290 8300 8310 8320 8330 8340 8350
8360 8370 8380 8390 8400 8410 8420
TGAGCCGAGTTTGCGCCACTGCACTCCAGCCTAGGTGACAGAGTGAGACTCCATCTCAAAAAAATAAAAATA
TGAGCCGAGTTTGCGCCACTGCACTCCAGCCTAGGTGACAGAGTGAGACTCCATCTCAAAAAAATAAAAATA
8360 8370 8380 8390 8400 8410 8420
8430 8440 8450 8460 8470 8480 8490
AAAATAAAAAAATGAAAAAAAAAAGAAAGTGAAGTATAGAGTATCTCATAGTTTGTCAGTGATAGAAACAGG
AAAATAAAAAAATGAAAAAAAAAAGAAAGTGAAGTATAGAGTATCTCATAGTTTGTCAGTGATAGAAACAGG
8430 8440 8450 8460 8470 8480 8490
8500 8510 8520 8530 8540 8550 8560
TTTCAAACTCAGTCAATCTGACCGTTTGATACATCTCAGACACCACTACATTCAGTAGTTTAGATGCCTAGA
TTTCAAACTCAGTCAATCTGACCGTTTGATACATCTCAGACACCACTACATTCAGTAGTTTAGATGCCTAGA
8500 8510 8520 8530 8540 8550 8560
8570 8580 8590 8600 8610 8620 8630 8640
ATAAATAGAGAAGGAAGGAGATGGCTCTTCTCTTGTCTCATTGTGTTTCTTCTGAGTGAGCTTGAATCACAT
ATAAATAGAGAAGGAAGGAGATGGCTCTTCTCTTGTCTCATTGTGTTTCTTCTGAGTGAGCTTGAATCACAT
8570 8580 8590 8600 8610 8620 8630 8640
8650 8660 8670 8680 8690 8700 8710
GAAGGGGAACAGCAGAAAACAACCAACTGATCCTCAGCTGTCATGTTTCCTTTAAAAGTCCCTGAAGGAAGG
GAAGGGGAACAGCAGAAAACAACCAACTGATCCTCAGCTGTCATGTTTCCTTTAAAAGTCCCTGAAGGAAGG
8650 8660 8670 8680 8690 8700 8710
8720 8730 8740 8750 8760 8770 8780
TCCTGGAATGTGACTCCCTTGCTCCTCTGTTGCTCTCTTTGGCATTCATTTCTTTGGACCCTACGCAAGGAC
TCCTGGAATGTGACTCCCTTGCTCCTCTGTTGCTCTCTTTGGCATTCATTTCTTTGGACCCTACGCAAGGAC
8720 8730 8740 8750 8760 8770 8780
8790 8800 8810 8820 8830 8840 8850
TGTAATTGGTGGGGACAGCTAGTGGCCCTGCTGGGCTTCACACACGGTGTCCTCCCTAGGCCAGTGCCTCTG
TGTAATTGGTGGGGACAGCTAGTGGCCCTGCTGGGCTTCACACACGGTGTCCTCCCTAGGCCAGTGCCTCTG
8790 8800 8810 8820 8830 8840 8850
8860 8870 8880 8890 8900 8910 8920
GAGTCAGAACTCTGGTGGTATTTCCCTCAATGAAGTGGAGTAAGCTCTCTCATTTTGAGATGGTATAATGGA
GAGTCAGAACTCTGGTGGTATTTCCCTCAATGAAGTGGAGTAAGCTCTCTCATTTTGAGATGGTATAATGGA
8860 8870 8880 8890 8900 8910 8920
8930 8940 8950 8960 8970 8980 8990 9000
AGCCACCAAGTGGCTTAGAGGATGCCCAGGTCCTTCCATGGAGCCACTGGGGTTCCGGTGCACATTAAAAAA
AGCCACCAAGTGGCTTAGAGGATGCCCAGGTCCTTCCATGGAGCCACTGGGGTTCCGGTGCACATTAAAAAA
8930 8940 8950 8960 8970 8980 8990 9000
9010 9020 9030 9040 9050 9060 9070
AAAATCTAACCAGGACATTCAGGAATTGCTAGATTCTGGGAAATCAGTTCACCATGTTCAAAAGAGTCTTTT
AAAATCTAACCAGGACATTCAGGAATTGCTAGATTCTGGGAAATCAGTTCACCATGTTCAAAAGAGTCTTTT
9010 9020 9030 9040 9050 9060 9070
9080 9090 9100 9110 9120 9130 9140
TTTTTTTTTTGAGACTCTATTGCCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGTAACCTCTGCCT
TTTTTTTTTTGAGACTCTATTGCCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGTAACCTCTGCCT
9080 9090 9100 9110 9120 9130 9140
9150 9160 9170 9180 9190 9200 9210
CCCAGGTTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCACCACCATGCCCG
CCCAGGTTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCACCACCATGCCCG
9150 9160 9170 9180 9190 9200 9210
9220 9230 9240 9250 9260 9270 9280
GCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCTCCTGACC
GCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCTCCTGACC
9220 9230 9240 9250 9260 9270 9280
9290 9300 9310 9320 9330 9340 9350 9360
TCGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGAGATTACAGGTGTGAGCCACCCTGCCCAGCCGTCAAA
TCGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGAGATTACAGGTGTGAGCCACCCTGCCCAGCCGTCAAA
9290 9300 9310 9320 9330 9340 9350 9360
9370 9380 9390 9400 9410 9420 9430
AGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAATG
AGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAATG
9370 9380 9390 9400 9410 9420 9430
9440 9450 9460 9470 9480 9490 9500
TGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCTTCTTTAG
TGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCTTCTTTAG
9440 9450 9460 9470 9480 9490 9500
9510 9520 9530 9540 9550 9560 9570
GCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTTTTTTGTGGTTAGAAAAGTTA
GCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTTTTTTGTGGTTAGAAAAGTTA
9510 9520 9530 9540 9550 9560 9570
9580 9590 9600 9610 9620 9630 9640
TGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTAT
TGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTAT
9580 9590 9600 9610 9620 9630 9640
9650 9660 9670 9680 9690 9700 9710 9720
TTAATTAGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCA
TTAATTAGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCA
9650 9660 9670 9680 9690 9700 9710 9720
9730 9740 9750 9760 9770 9780 9790
TATACTTTAATAAATGTATATTGTATTGTATACTGCATGATTTTATTGAAGTTCTTGTTCATCTTGTGTATA
TATACTTTAATAAATGTATATTGTATTGTATACTGCATGATTTTATTGAAGTTCTTGTTCATCTTGTGTATA
9730 9740 9750 9760 9770 9780 9790
9800 9810 9820 9830 9840 9850 9860
TACTTAATCGCTTTGTCATTTTGGAGACATTTATTTTGCTTCTAATTTCTTTACATTTTGTCTTACGGAATA
TACTTAATCGCTTTGTCATTTTGGAGACATTTATTTTGCTTCTAATTTCTTTACATTTTGTCTTACGGAATA
9800 9810 9820 9830 9840 9850 9860
9870 9880 9890 9900 9910 9920 9930
TTTTCATTCAACTGTGGTAGCCGAATTAATCGTGTTTCTTCACTCTAGGGACATTGTCGTCTAAGTTGTAAG
TTTTCATTCAACTGTGGTAGCCGAATTAATCGTGTTTCTTCACTCTAGGGACATTGTCGTCTAAGTTGTAAG
9870 9880 9890 9900 9910 9920 9930
9940 9950 9960 9970 9980 9990 10000
ACATTGGTTATTTTACCAGCAAACCATTCTGAAAGCATATGACAAATTATTTCTCTCTTAATATCTTACTAT
ACATTGGTTATTTTACCAGCAAACCATTCTGAAAGCATATGACAAATTATTTCTCTCTTAATATCTTACTAT
9940 9950 9960 9970 9980 9990 10000
10010 10020 10030 10040 10050 10060 10070 10080
ACTGAAAGCAGACTGCTATAAGGCTTCACTTACTCTTCTACCTCATAAGGAATATGTTACAATTAATTTATT

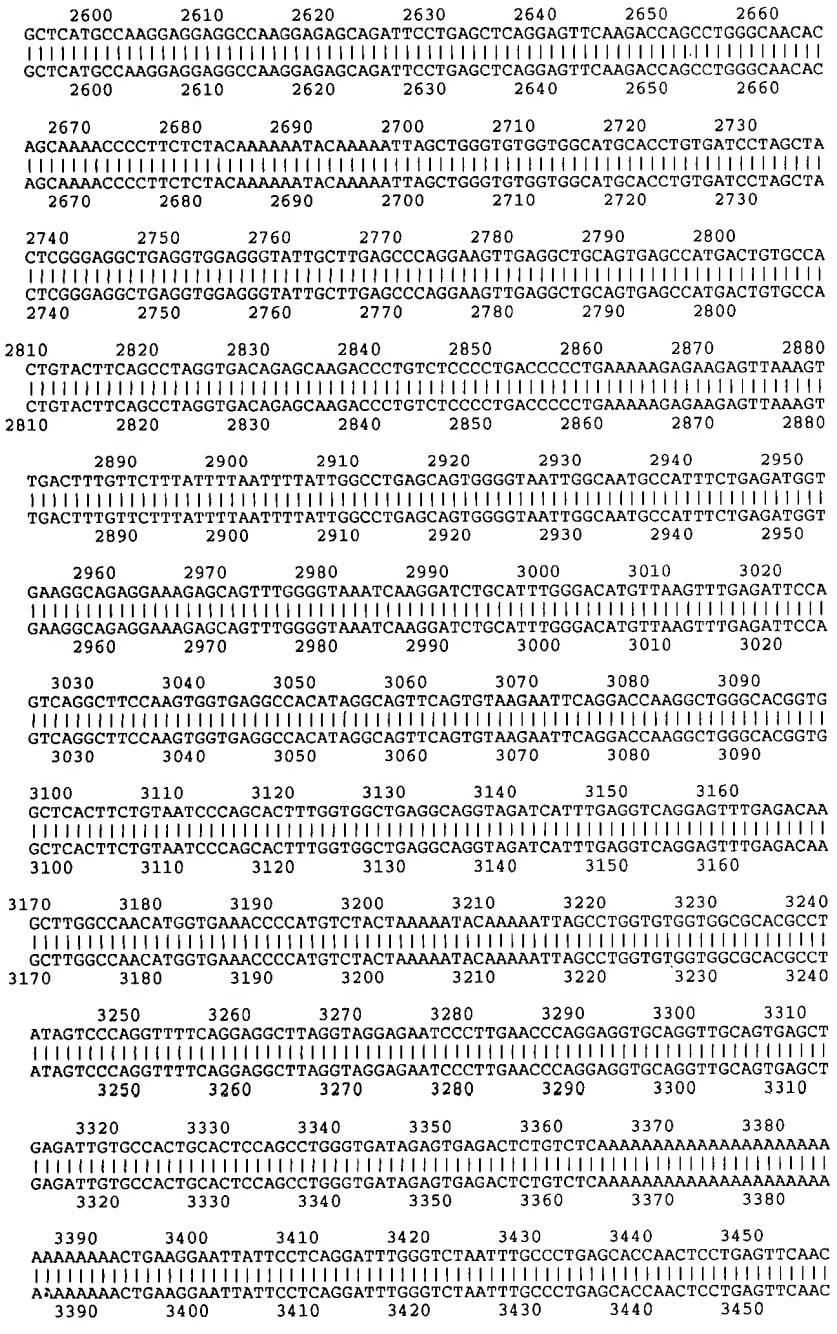


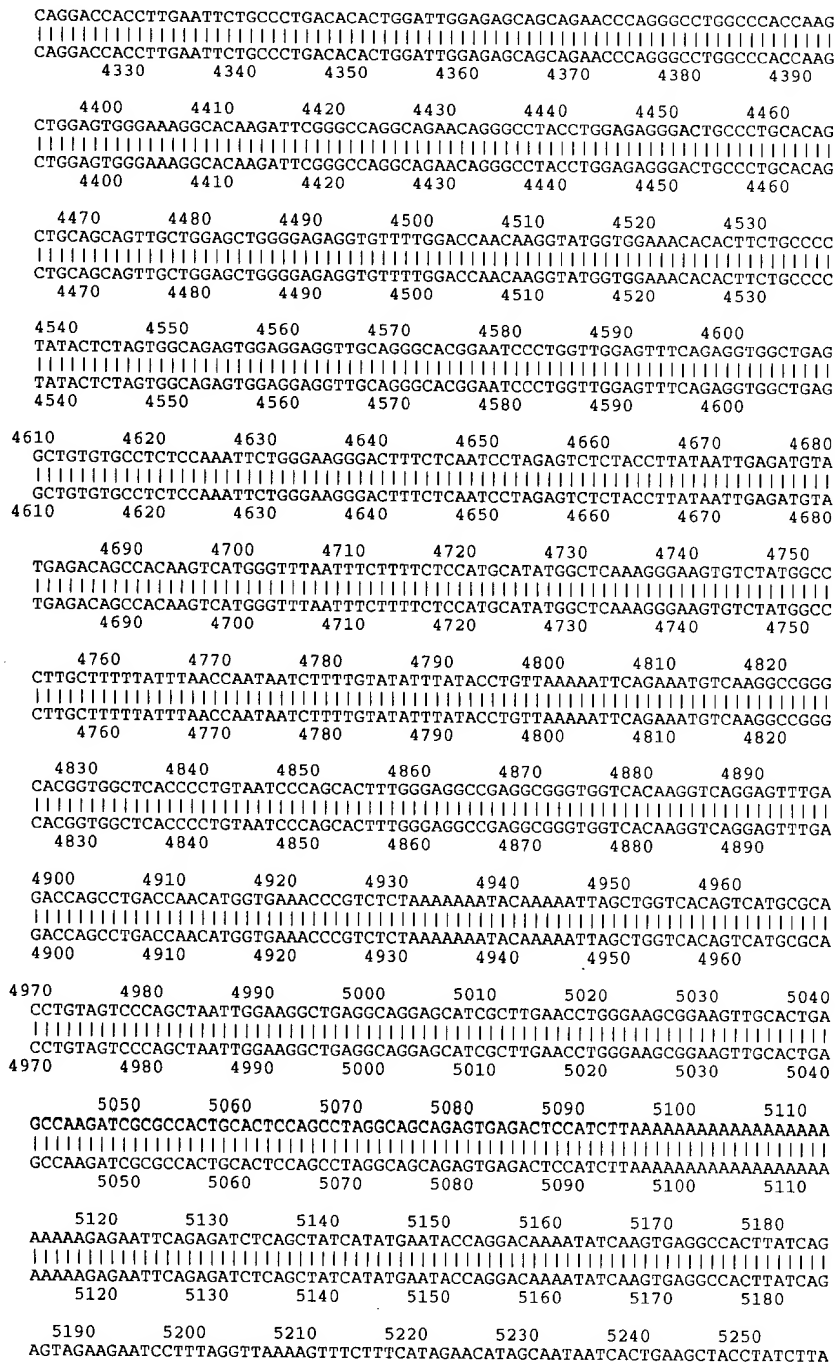
4. US,09-497-957-1 (1-10825)
US-09-497-957-1 Sequence 1, Application US/09497957

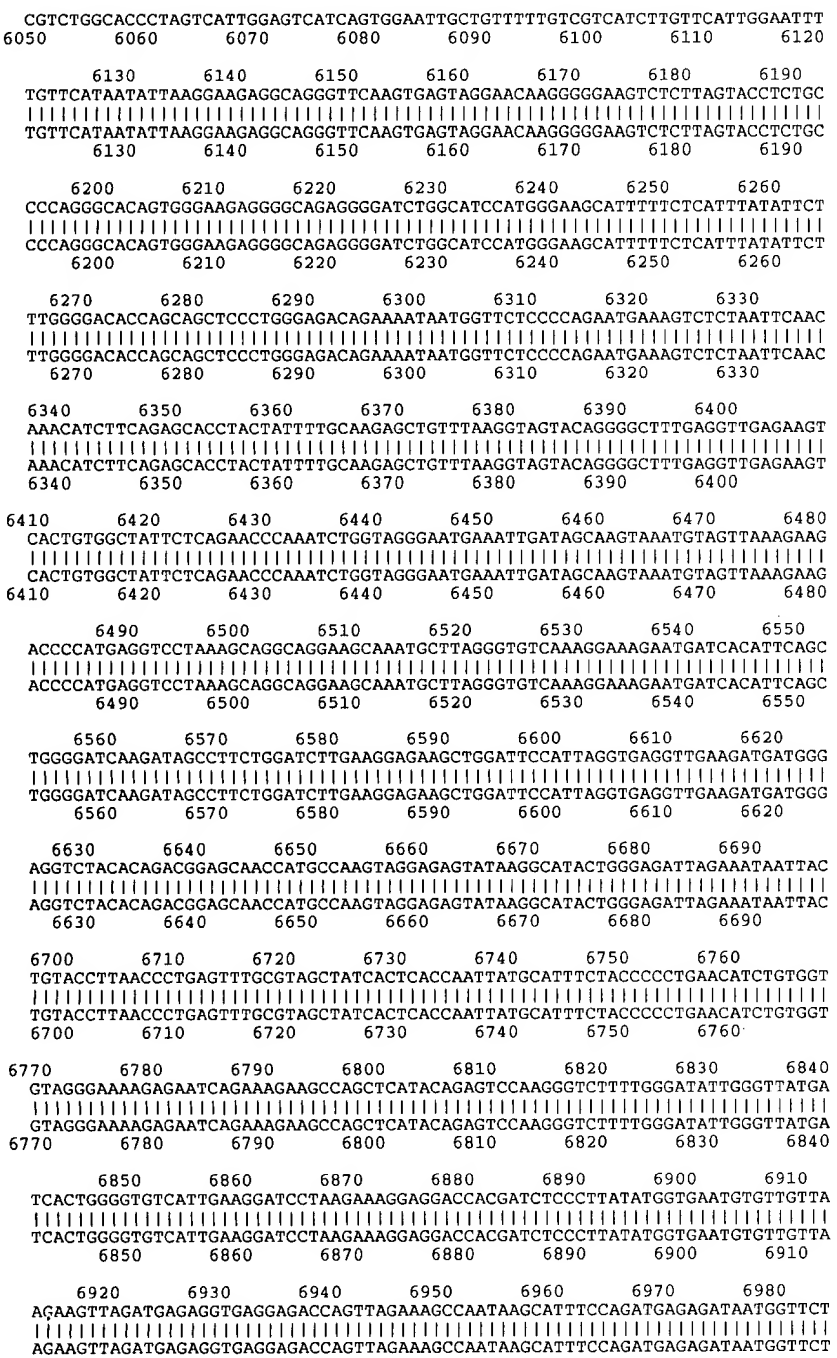


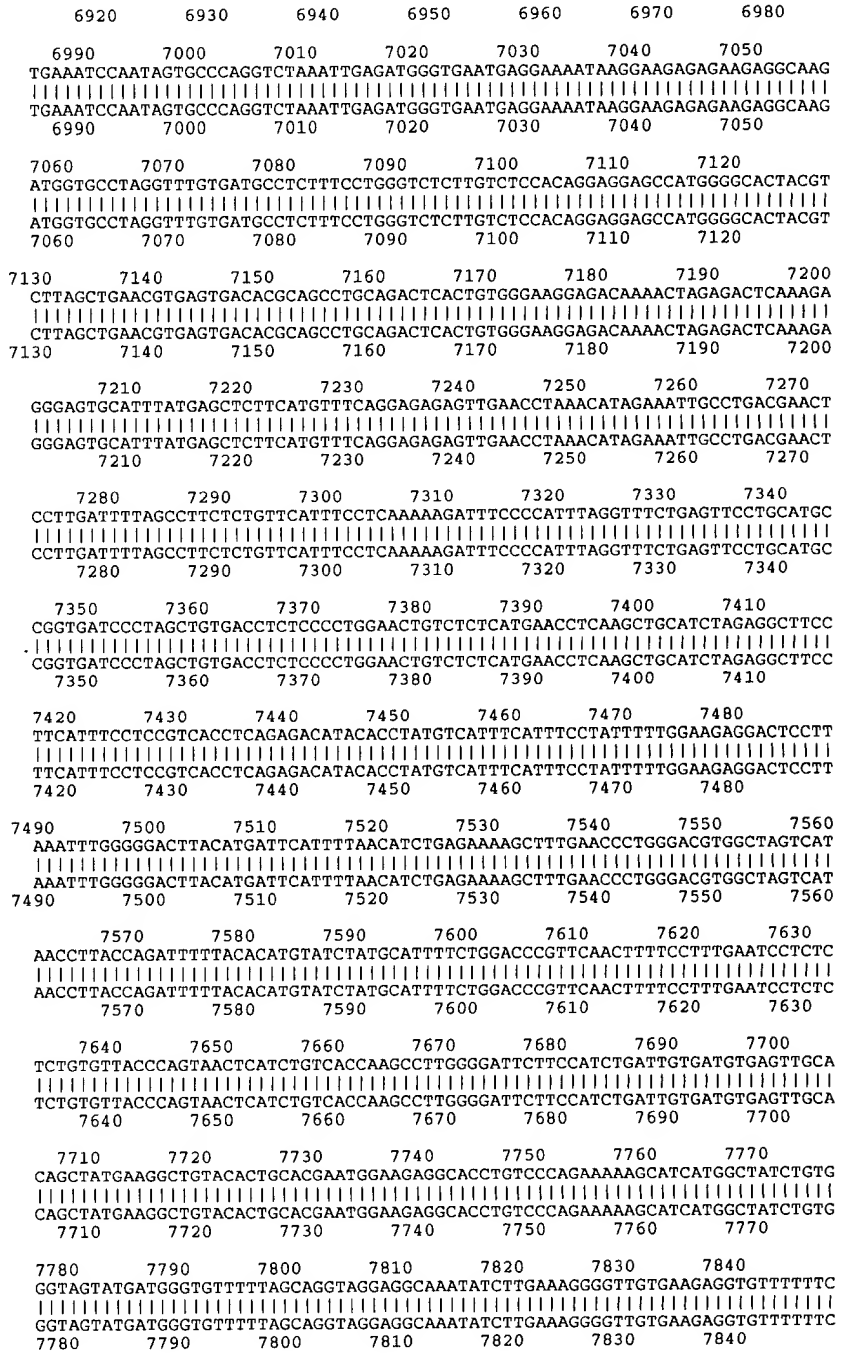

TTTCCAATGTCAGCTGTGCAGTTTTTTCCCCAGTCATCTCCAAACAGGAAGTTCTTCCCTGAGTGCTTGCCG
800 810 820 830 840 850 860
870 880 890 900 910 920 930
AGAAGGCTGAGCAAACCCACAGCAGGATCCGCACGGGGTTTCCACCTCAGAACGAATGCGTTGGGCGGTGGG
AGAAGGCTGAGCAAACCCACAGCAGGATCCGCACGGGGTTTCCACCTCAGAACGAATGCGTTGGGCGGTGGG
870 880 890 900 910 920 930
940 950 960 970 980 990 1000
GGCGCGAAAGAGTGGCGTTGGGGATCTGAATTCTTCACCATTCCACCCACTTTTGGTGAGACCTGGGGTGGA
GGCGCGAAAGAGTGGCGTTGGGGATCTGAATTCTTCACCATTCCACCCACTTTTGGTGAGACCTGGGGTGGA
940 950 960 970 980 990 1000
1010 1020 1030 1040 1050 1060 1070 1080
GGTCTCTAGGGTGGGAGGCTCCTGAGAGAGGCCTACCTCGGGCCTTTCCCCACTCTTGGCAATTGTTCTTTT
GGTCTCTAGGGTGGGAGGCTCCTGAGAGAGGCCTACCTCGGGCCTTTCCCCACTCTTGGCAATTGTTCTTTT
1010 1020 1030 1040 1050 1060 1070 1080
1090 1100 1110 1120 1130 1140 1150
GCCTGGAAAATTAAGTATATGTTAGTTTTGAACGTTTGAACTGAACAATTCTCTTTTCGGCTAGGCTTTATT
GCCTGGAAAATTAAGTATATGTTAGTTTTGAACGTTTGAACTGAACAATTCTCTTTTCGGCTAGGCTTTATT
1090 1100 1110 1120 1130 1140 1150
1160 1170 1180 1190 1200 1210 1220
GATTTGCAATGTGCTGTGTAATTAAGAGGCCTCTCTACAAAGTACTGATAATGAACATGTAAGCAATGCACT
GATTTGCAATGTGCTGTGTAATTAAGAGGCCTCTCTACAAAGTACTGATAATGAACATGTAAGCAATGCACT
1160 1170 1180 1190 1200 1210 1220
1230 1240 1250 1260 1270 1280 1290
CACTTCTAAGTTACATTCATATCTGATCTTATTTGATTTTCACTAGGCATAGGGAGGTAGGAGCTAATAATA
CACTTCTAAGTTACATTCATATCTGATCTTATTTGATTTTCACTAGGCATAGGGAGGTAGGAGCTAATAATA
1230 1240 1250 1260 1270 1280 1290
1300 1310 1320 1330 1340 1350 1360
CGTTTATTTTACTAGAAGTTAACTGGAATTCAGATTATATAACTCTTTTCAGGTTACAAAGAACATAAATAA
CGTTTATTTTACTAGAAGTTAACTGGAATTCAGATTATATAACTCTTTTCAGGTTACAAAGAACATAAATAA
1300 1310 1320 1330 1340 1350 1360
1370 1380 1390 1400 1410 1420 1430 1440
TCTGGTTTTCTGATGTTATTTCAAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTGATTTTGCCCTGTAG
TCTGGTTTTCTGATGTTATTTCAAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTGATTTTGCCCTGTAG
1370 1380 1390 1400 1410 1420 1430 1440
1450 1460 1470 1480 1490 1500 1510
TGTAGCACAGTGTTCTGTGGGTCACACGCCGGCCTCAGCACAGCACTTTGAGTTTTGGTACTACGTGTATCC
TGTAGCACAGTGTTCTGTGGGTCACACGCCGGCCTCAGCACAGCACTTTGAGTTTTGGTACTACGTGTATCC
1450 1460 1470 1480 1490 1500 1510
1520 1530 1540 1550 1560 1570 1580
ACATTTTACACATGACAAGAATGAGGCATGGCACGGCCTGCTTCCTGGCAAATTTATTCAATGGTACACTGG
ACATTTTACACATGACAAGAATGAGGCATGGCACGGCCTGCTTCCTGGCAAATTTATTCAATGGTACACTGG
1520 1530 1540 1550 1560 1570 1580
1590 1600 1610 1620 1630 1640 1650
GCTTTGGTGGCAGAGCTCATGTCTCCACTTCATAGCTATGATTCTTAAACATCACACTGCATTAGAGGTTGA
GCTTTGGTGGCAGAGCTCATGTCTCCACTTCATAGCTATGATTCTTAAACATCACACTGCATTAGAGGTTGA
1590 1600 1610 1620 1630 1640 1650
1660 1670 1680 1690 1700 1710 1720
ATAATAAAATTTCATGTTGAGCAGAAATATTCATTGTTTACAAGTGTAAATGAGTCCCAGCCATGTGTTGCA
ATAATAAAATTTCATGTTGAGCAGAAATATTCATTGTTTACAAGTGTAAATGAGTCCCAGCCATGTGTTGCA

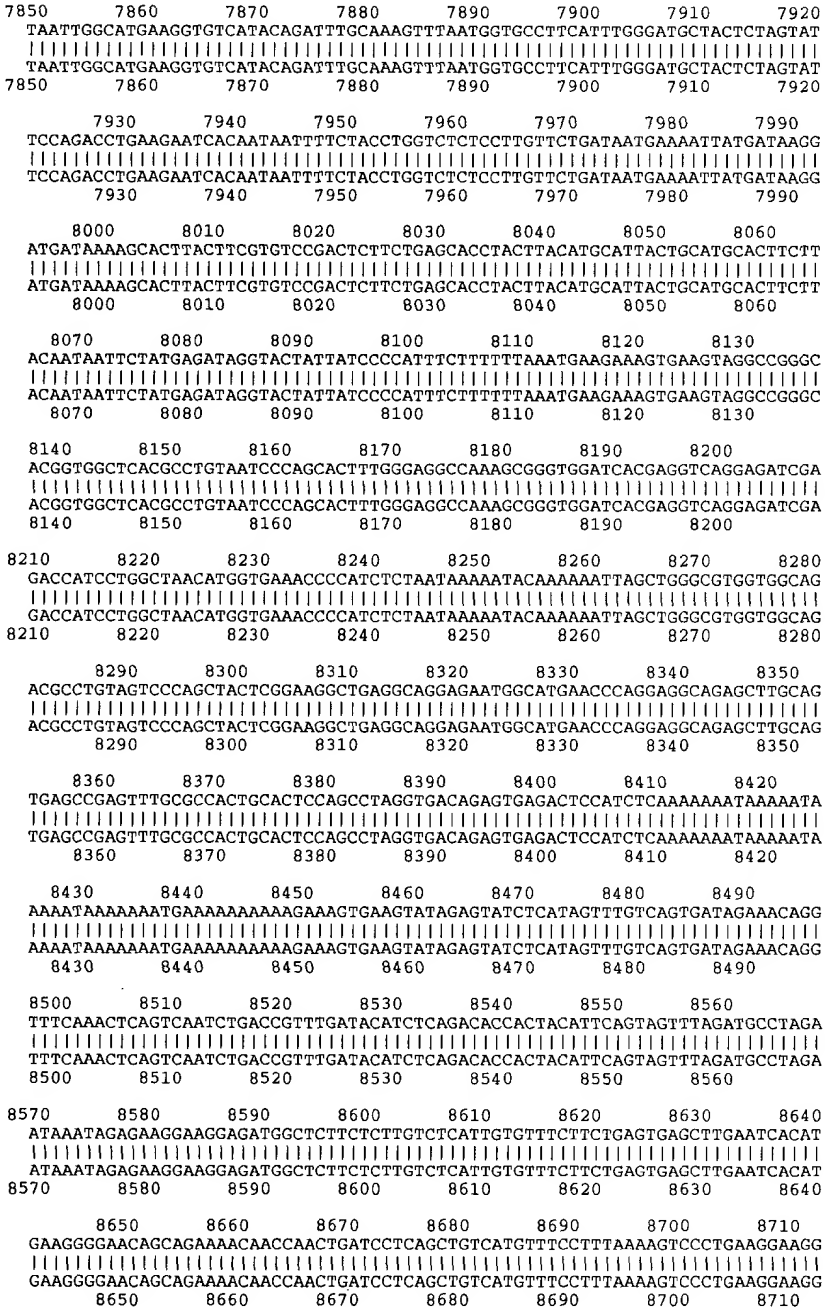

1660 1670 1680 1690 1700 1710 1720
1730 1740 1750 1760 1770 1780 1790 1800
CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
CTGTTCAAGCCCCAAGGGAGAGAGCAGGGAAACAAGTCTTTACCCTTTGATATTTTGCATTCTAGTGGGAGA
1730 1740 1750 1760 1770 1780 1790 1800
1810 1820 1830 1840 1850 1860 1870
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
GATGACAATAAGCAAATGAGCAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGAAG
1810 1820 1830 1840 1850 1860 1870
1880 1890 1900 1910 1920 1930 1940
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
TCAGGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGGAAATAAGAATGATATTGACTGGGAG
1880 1890 1900 1910 1920 1930 1940
1950 1960 1970 1980 1990 2000 2010
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
CAGTATTTCCCAGGCAAACTGAGTGGGCCTGGCAAGTTGGATTAAAAAGCGGGTTTTCTCAGCACTACTCAT
1950 1960 1970 1980 1990 2000 2010
2020 2030 2040 2050 2060 2070 2080
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
GTGTGTGTGTGTGGGGGGGGGGGGCGGCGTGGGGGTGGGAAGGGGGACTACCATCTGCATGTAGGATGTCTA
2020 2030 2040 2050 2060 2070 2080
2090 2100 2110 2120 2130 2140 2150 2160
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAATGTCTCT
GCAGTATCCTGTCCTCCCTACTCACTAGGTGCTAGGAGCACTCCCCCAGTCTTGACAACCAAAAATGTCTCT
2090 2100 2110 2120 2130 2140 2150 2160
2170 2180 2190 2200 2210 2220 2230
AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA
AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTGGTTAAGAAGCTCGGGTTGAAAAAAATAAACAAGTA
2170 2180 2190 2200 2210 2220 2230
2240 2250 2260 2270 2280 2290 2300
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
GTGCTGGGGAGTAGAGGCCAAGAAGTAGGTAATGGGCTCAGAAGAGGAGCCACAAACAAGGTTGTGCAGGCG
2240 2250 2260 2270 2280 2290 2300
2310 2320 2330 2340 2350 2360 2370
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
CCTGTAGGCTGTGGTGTGAATTCTAGCCAAGGAGTAACAGTGATCTGTCACAGGCTTTTAAAAGATTGCTCT
2310 2320 2330 2340 2350 2360 2370
2380 2390 2400 2410 2420 2430 2440
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
GGCTGCTATGTGGAAAGCAGAATGAAGGGAGCAACAGTAAAAGCAGGGAGCCCAGCCAGGAAGCTGTTACAC
2380 2390 2400 2410 2420 2430 2440
2450 2460 2470 2480 2490 2500 2510 2520
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
AGTCCAGGCAAGAGGTAGTGGAGTGGGCTGGGTGGGAACAGAAAAGGGAGTGACAAACCATTGTCTCCTGAA
2450 2460 2470 2480 2490 2500 2510 2520
2530 2540 2550 2560 2570 2580 2590
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
TATATTCTGAAGGAAGTTGCTGAAGGATTCTATGTTGTGTGAGAGAAAGAGAAGAATTGGCTGGGTGTAGTA
2530 2540 2550 2560 2570 2580 2590

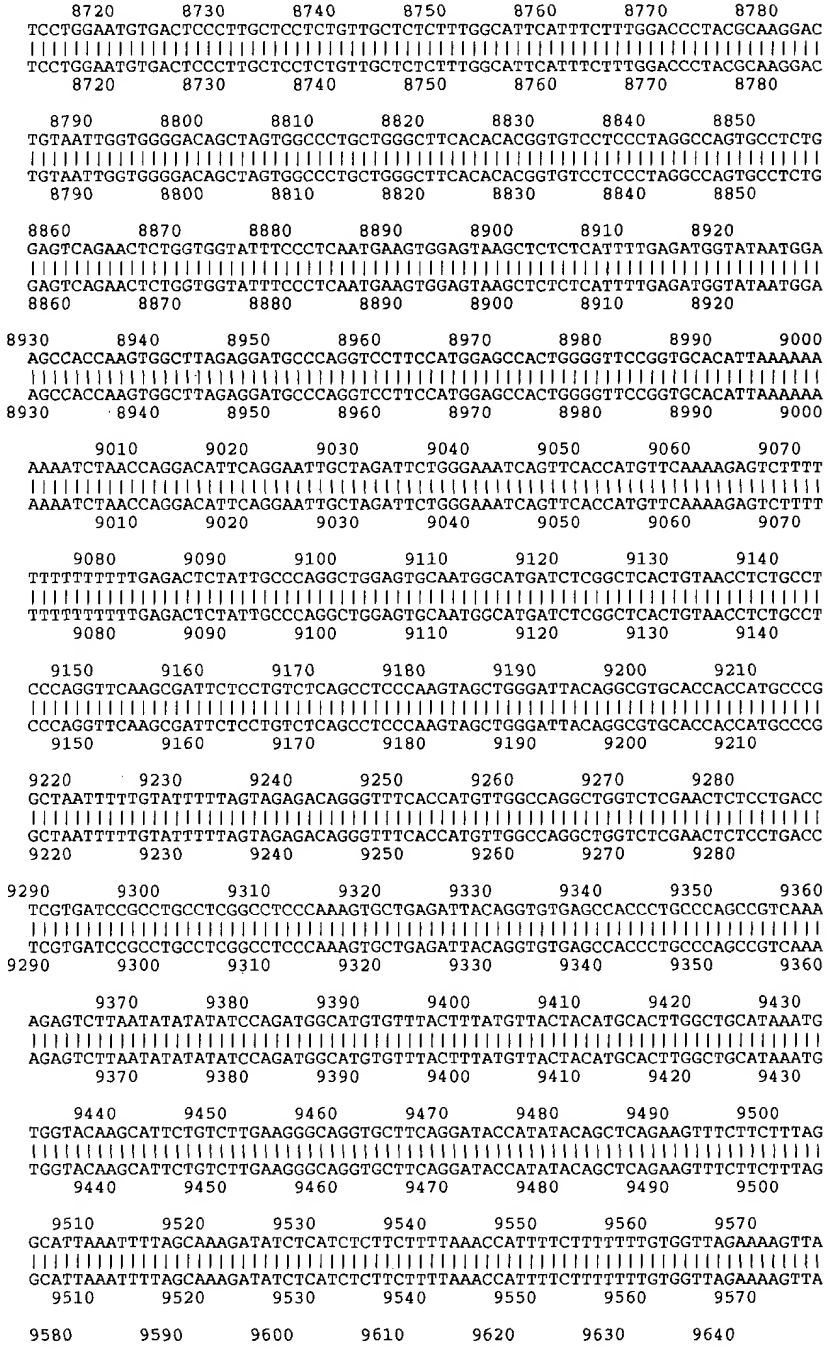

2600 2610 2620 2630 2640 2650 2660
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
GCTCATGCCAAGGAGGAGGCCAAGGAGAGCAGATTCCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACAC
2600 2610 2620 2630 2640 2650 2660
2670 2680 2690 2700 2710 2720 2730
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
AGCAAAACCCCTTCTCTACAAAAAATACAAAAATTAGCTGGGTGTGGTGGCATGCACCTGTGATCCTAGCTA
2670 2680 2690 2700 2710 2720 2730
2740 2750 2760 2770 2780 2790 2800
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
CTCGGGAGGCTGAGGTGGAGGGTATTGCTTGAGCCCAGGAAGTTGAGGCTGCAGTGAGCCATGACTGTGCCA
2740 2750 2760 2770 2780 2790 2800
2810 2820 2830 2840 2850 2860 2870 2880
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
CTGTACTTCAGCCTAGGTGACAGAGCAAGACCCTGTCTCCCCTGACCCCCTGAAAAAGAGAAGAGTTAAAGT
2810 2820 2830 2840 2850 2860 2870 2880
2890 2900 2910 2920 2930 2940 2950
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
TGACTTTGTTCTTTATTTTAATTTTATTGGCCTGAGCAGTGGGGTAATTGGCAATGCCATTTCTGAGATGGT
2890 2900 2910 2920 2930 2940 2950
2960 2970 2980 2990 3000 3010 3020
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
GAAGGCAGAGGAAAGAGCAGTTTGGGGTAAATCAAGGATCTGCATTTGGGACATGTTAAGTTTGAGATTCCA
2960 2970 2980 2990 3000 3010 3020
3030 3040 3050 3060 3070 3080 3090
GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
GTCAGGCTTCCAAGTGGTGAGGCCACATAGGCAGTTCAGTGTAAGAATTCAGGACCAAGGCTGGGCACGGTG
3030 3040 3050 3060 3070 3080 3090
3100 3110 3120 3130 3140 3150 3160
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
GCTCACTTCTGTAATCCCAGCACTTTGGTGGCTGAGGCAGGTAGATCATTTGAGGTCAGGAGTTTGAGACAA
3100 3110 3120 3130 3140 3150 3160
3170 3180 3190 3200 3210 3220 3230 3240
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
GCTTGGCCAACATGGTGAAACCCCATGTCTACTAAAAATACAAAAATTAGCCTGGTGTGGTGGCGCACGCCT
3170 3180 3190 3200 3210 3220 3230 3240
3250 3260 3270 3280 3290 3300 3310
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
ATAGTCCCAGGTTTTCAGGAGGCTTAGGTAGGAGAATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAGCT
3250 3260 3270 3280 3290 3300 3310
3320 3330 3340 3350 3360 3370 3380
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
GAGATTGTGCCACTGCACTCCAGCCTGGGTGATAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA
3320 3330 3340 3350 3360 3370 3380
3390 3400 3410 3420 3430 3440 3450
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
AAAAAAAACTGAAGGAATTATTCCTCAGGATTTGGGTCTAATTTGCCCTGAGCACCAACTCCTGAGTTCAAC
3390 3400 3410 3420 3430 3440 3450
3460 3470 3480 3490 3500 3510 3520
TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA
TACCATGGCTAGACACACCTTAACATTTTCTAGAATCCACCAGCTTTAGTGGAGTCTGTCTAATCATGAGTA
3460 3470 3480 3490 3500 3510 3520
3530 3540 3550 3560 3570 3580 3590 3600
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
TTGGAATAGGATCTGGGGGCAGTGAGGGGGTGGCAGCCACGTGTGGCAGAGAAAAGCACACAAGGAAAGAGC
3530 3540 3550 3560 3570 3580 3590 3600
3610 3620 3630 3640 3650 3660 3670
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
ACCCAGGACTGTCATATGGAAGAAAGACAGGACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCT
3610 3620 3630 3640 3650 3660 3670
3680 3690 3700 3710 3720 3730 3740
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
GATGGTATGAGTTGATGCAGGTGTGTGGAGCCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCC
3680 3690 3700 3710 3720 3730 3740
3750 3760 3770 3780 3790 3800 3810
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
TGTTGCTCTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCT
3750 3760 3770 3780 3790 3800 3810
3820 3830 3840 3850 3860 3870 3880
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATNATGAGAGTCGCCGTGT
TTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGT
3820 3830 3840 3850 3860 3870 3880
3890 3900 3910 3920 3930 3940 3950 3960
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
GGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA
3890 3900 3910 3920 3930 3940 3950 3960
3970 3980 3990 4000 4010 4020 4030
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
AGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGGTATGTG
3970 3980 3990 4000 4010 4020 4030
4040 4050 4060 4070 4080 4090 4100
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
GAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAAACAGCTGG
4040 4050 4060 4070 4080 4090 4100
4110 4120 4130 4140 4150 4160 4170
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
AAGTCTGAGGTCTTGTGGGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTGGTCCTTGGGGAT
4110 4120 4130 4140 4150 4160 4170
4180 4190 4200 4210 4220 4230 4240
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
GGTGGAAATAGGGACCTATTCCTTTGGTTGCAGTTAACAAGGCTGGGGATTTTTCCAGAGTCCCACACCCTG
4180 4190 4200 4210 4220 4230 4240
4250 4260 4270 4280 4290 4300 4310 4320
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
CAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGG
4250 4260 4270 4280 4290 4300 4310 4320
4330 4340 4350 4360 4370 4380 4390

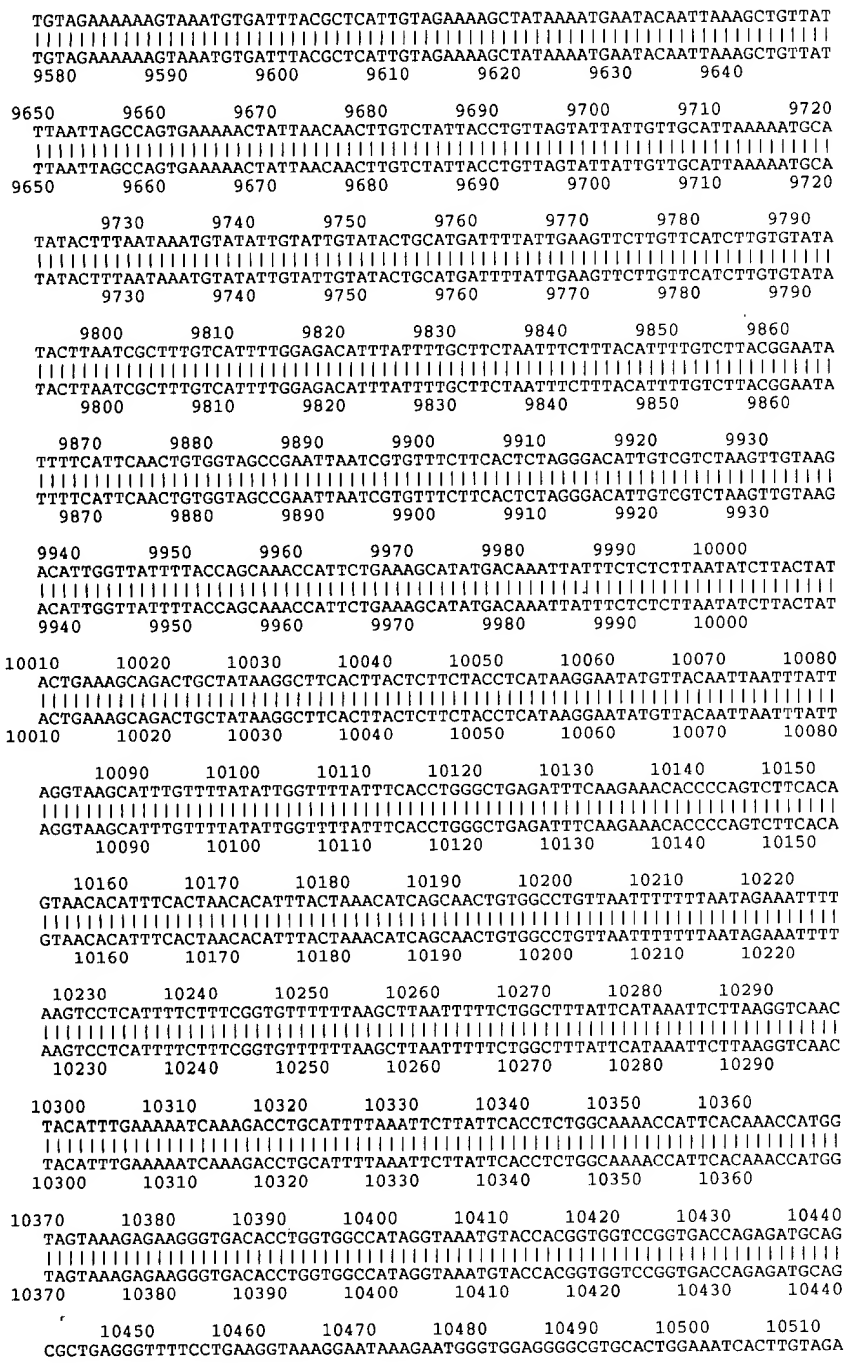

CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
CAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAG
4330 4340 4350 4360 4370 4380 4390
4400 4410 4420 4430 4440 4450 4460
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
CTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAG
4400 4410 4420 4430 4440 4450 4460
4470 4480 4490 4500 4510 4520 4530
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
CTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGGTATGGTGGAAACACACTTCTGCCCC
4470 4480 4490 4500 4510 4520 4530
4540 4550 4560 4570 4580 4590 4600
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
TATACTCTAGTGGCAGAGTGGAGGAGGTTGCAGGGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAG
4540 4550 4560 4570 4580 4590 4600
4610 4620 4630 4640 4650 4660 4670 4680
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
GCTGTGTGCCTCTCCAAATTCTGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTA
4610 4620 4630 4640 4650 4660 4670 4680
4690 4700 4710 4720 4730 4740 4750
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
TGAGACAGCCACAAGTCATGGGTTTAATTTCTTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATGGCC
4690 4700 4710 4720 4730 4740 4750
4760 4770 4780 4790 4800 4810 4820
CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGG
CTTGCTTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAAGGCCGGG
4760 4770 4780 4790 4800 4810 4820
4830 4840 4850 4860 4870 4880 4890
CACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGTCACAAGGTCAGGAGTTTGA
CACGGTGGCTCACCCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGTCACAAGGTCAGGAGTTTGA
4830 4840 4850 4860 4870 4880 4890
4900 4910 4920 4930 4940 4950 4960
GACCAGCCTGACCAACATGGTGAAACCCGTCTCTAAAAAAATACAAAAATTAGCTGGTCACAGTCATGCGCA
GACCAGCCTGACCAACATGGTGAAACCCGTCTCTAAAAAAATACAAAAATTAGCTGGTCACAGTCATGCGCA
4900 4910 4920 4930 4940 4950 4960
4970 4980 4990 5000 5010 5020 5030 5040
CCTGTAGTCCCAGCTAATTGGAAGGCTGAGGCAGGAGCATCGCTTGAACCTGGGAAGCGGAAGTTGCACTGA
CCTGTAGTCCCAGCTAATTGGAAGGCTGAGGCAGGAGCATCGCTTGAACCTGGGAAGCGGAAGTTGCACTGA
4970 4980 4990 5000 5010 5020 5030 5040
5050 5060 5070 5080 5090 5100 5110
GCCAAGATCGCGCCACTGCACTCCAGCCTAGGCAGCAGAGTGAGACTCCATCTTAAAAAAAAAAAAAAAAAA
GCCAAGATCGCGCCACTGCACTCCAGCCTAGGCAGCAGAGTGAGACTCCATCTTAAAAAAAAAAAAAAAAAA
5050 5060 5070 5080 5090 5100 5110
5120 5130 5140 5150 5160 5170 5180
AAAAAGAGAATTCAGAGATCTCAGCTATCATATGAATACCAGGACAAAATATCAAGTGAGGCCACTTATCAG
AAAAAGAGAATTCAGAGATCTCAGCTATCATATGAATACCAGGACAAAATATCAAGTGAGGCCACTTATCAG
5120 5130 5140 5150 5160 5170 5180
5190 5200 5210 5220 5230 5240 5250
AGTAGAAGAATCCTTTAGGTTAAAAGTTTCTTTCATAGAACATAGCAATAATCACTGAAGCTACCTATCTTA
AGTAGAAGAATCCTTTAGGTTAAAAGTTTCTTTCATAGAACATAGCAATAATCACTGAAGCTACCTATCTTA
5190 5200 5210 5220 5230 5240 5250
5260 5270 5280 5290 5300 5310 5320
CAAGTCCGCTTCTTATAACAATGCCTCCTAGGTTGACCCAGGTGAAACTGACCATCTGTATTCAATCATTTT
CAAGTCCGCTTCTTATAACAATGCCTCCTAGGTTGACCCAGGTGAAACTGACCATCTGTATTCAATCATTTT
5260 5270 5280 5290 5300 5310 5320
5330 5340 5350 5360 5370 5380 5390 5400
CAATGCACATAAAGGGCAATTTTATCTATCAGAACAAAGAACATGGGTAACAGATATGTATATTTACATGTG
CAATGCACATAAAGGGCAATTTTATCTATCAGAACAAAGAACATGGGTAACAGATATGTATATTTACATGTG
5330 5340 5350 5360 5370 5380 5390 5400
5410 5420 5430 5440 5450 5460 5470
AGGAGAACAAGCTGATCTGACTGCTCTCCAAGTGACACTGTGTTAGAGTCCAATCTTAGGACACAAAATGGT
AGGAGAACAAGCTGATCTGACTGCTCTCCAAGTGACACTGTGTTAGAGTCCAATCTTAGGACACAAAATGGT
5410 5420 5430 5440 5450 5460 5470
5480 5490 5500 5510 5520 5530 5540
GTCTCTCCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTT
GTCTCTCCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTT
5480 5490 5500 5510 5520 5530 5540
5550 5560 5570 5580 5590 5600 5610
CCAGTCTTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTG
CCAGTCTTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTG
5550 5560 5570 5580 5590 5600 5610
5620 5630 5640 5650 5660 5670 5680
GTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG
GTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG
5620 5630 5640 5650 5660 5670 5680
5690 5700 5710 5720 5730 5740 5750 5760
AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTG
AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTG
5690 5700 5710 5720 5730 5740 5750 5760
5770 5780 5790 5800 5810 5820 5830
CCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACG
CCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACG
5770 5780 5790 5800 5810 5820 5830
5840 5850 5860 5870 5880 5890 5900
TNCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGG
TGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGG
5840 5850 5860 5870 5880 5890 5900
5910 5920 5930 5940 5950 5960 5970
AGCTGAGAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGC
AGCTGAGAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGC
5910 5920 5930 5940 5950 5960 5970
5980 5990 6000 6010 6020 6030 6040
TCTTTGTTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCAC
TCTTTGTTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCAC
5980 5990 6000 6010 6020 6030 6040
6050 6060 6070 6080 6090 6100 6110 6120
CGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTT

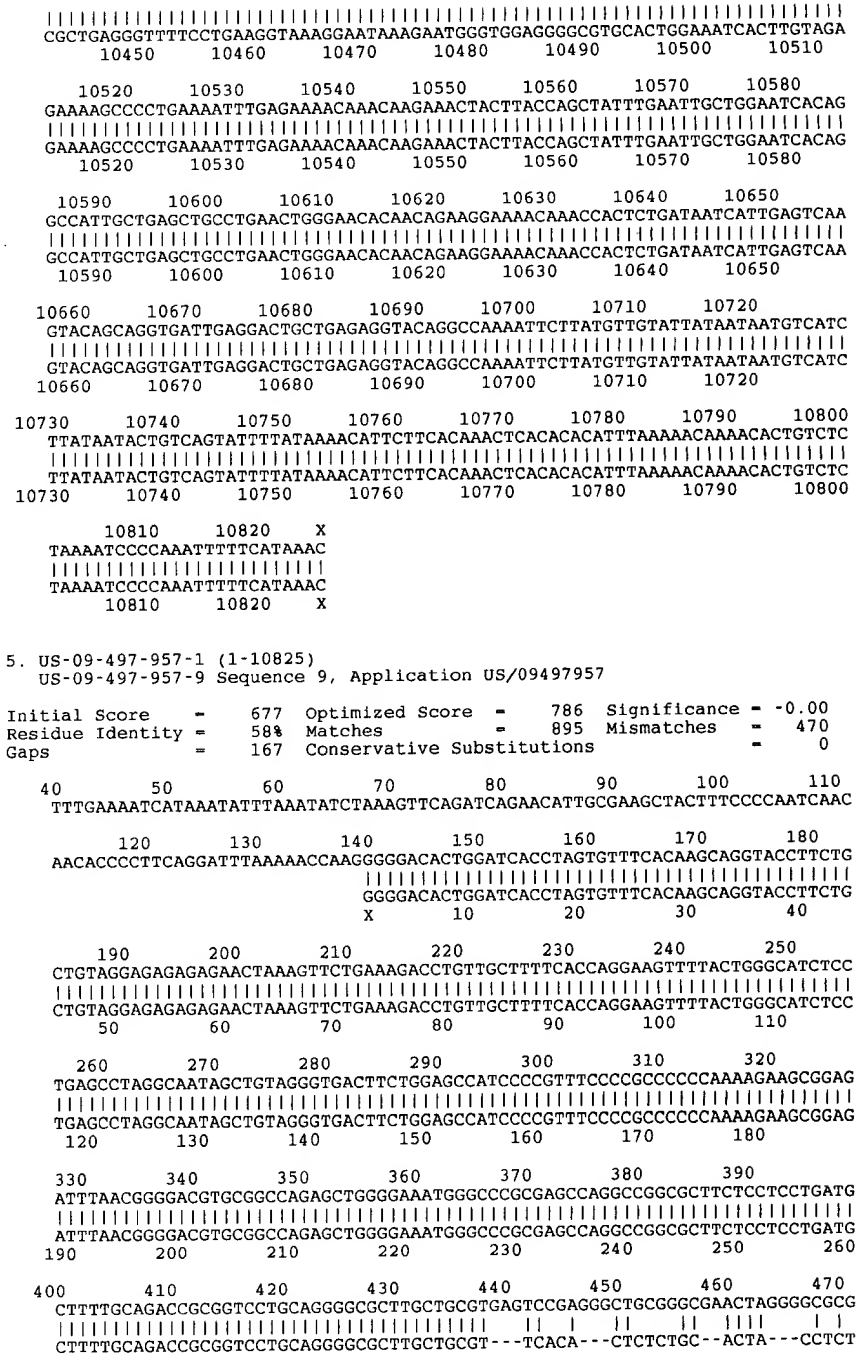

CGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTT
6050 6060 6070 6080 6090 6100 6110 6120
6130 6140 6150 6160 6170 6180 6190
TGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGTGAGTAGGAACAAGGGGGAAGTCTCTTAGTACCTCTGC
TGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGTGAGTAGGAACAAGGGGGAAGTCTCTTAGTACCTCTGC
6130 6140 6150 6160 6170 6180 6190
6200 6210 6220 6230 6240 6250 6260
CCCAGGGCACAGTGGGAAGAGGGGCAGAGGGGATCTGGCATCCATGGGAAGCATTTTTCTCATTTATATTCT
CCCAGGGCACAGTGGGAAGAGGGGCAGAGGGGATCTGGCATCCATGGGAAGCATTTTTCTCATTTATATTCT
6200 6210 6220 6230 6240 6250 6260
6270 6280 6290 6300 6310 6320 6330
TTGGGGACACCAGCAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCCAGAATGAAAGTCTCTAATTCAAC
TTGGGGACACCAGCAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCCAGAATGAAAGTCTCTAATTCAAC
6270 6280 6290 6300 6310 6320 6330
6340 6350 6360 6370 6380 6390 6400
AAACATCTTCAGAGCACCTACTATTTTGCAAGAGCTGTTTAAGGTAGTACAGGGGCTTTGAGGTTGAGAAGT
AAACATCTTCAGAGCACCTACTATTTTGCAAGAGCTGTTTAAGGTAGTACAGGGGCTTTGAGGTTGAGAAGT
6340 6350 6360 6370 6380 6390 6400
6410 6420 6430 6440 6450 6460 6470 6480
CACTGTGGCTATTCTCAGAACCCAAATCTGGTAGGGAATGAAATTGATAGCAAGTAAATGTAGTTAAAGAAG
CACTGTGGCTATTCTCAGAACCCAAATCTGGTAGGGAATGAAATTGATAGCAAGTAAATGTAGTTAAAGAAG
6410 6420 6430 6440 6450 6460 6470 6480
6490 6500 6510 6520 6530 6540 6550
ACCCCATGAGGTCCTAAAGCAGGCAGGAAGCAAATGCTTAGGGTGTCAAAGGAAAGAATGATCACATTCAGC
ACCCCATGAGGTCCTAAAGCAGGCAGGAAGCAAATGCTTAGGGTGTCAAAGGAAAGAATGATCACATTCAGC
6490 6500 6510 6520 6530 6540 6550
6560 6570 6580 6590 6600 6610 6620
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
TGGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAAGCTGGATTCCATTAGGTGAGGTTGAAGATGATGGG
6560 6570 6580 6590 6600 6610 6620
6630 6640 6650 6660 6670 6680 6690
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
AGGTCTACACAGACGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATACTGGGAGATTAGAAATAATTAC
6630 6640 6650 6660 6670 6680 6690
6700 6710 6720 6730 6740 6750 6760
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
TGTACCTTAACCCTGAGTTTGCGTAGCTATCACTCACCAATTATGCATTTCTACCCCCTGAACATCTGTGGT
6700 6710 6720 6730 6740 6750 6760
6770 6780 6790 6800 6810 6820 6830 6840
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
GTAGGGAAAAGAGAATCAGAAAGAAGCCAGCTCATACAGAGTCCAAGGGTCTTTTGGGATATTGGGTTATGA
6770 6780 6790 6800 6810 6820 6830 6840
6850 6860 6870 6880 6890 6900 6910
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
TCACTGGGGTGTCATTGAAGGATCCTAAGAAAGGAGGACCACGATCTCCCTTATATGGTGAATGTGTTGTTA
6850 6860 6870 6880 6890 6900 6910
6920 6930 6940 6950 6960 6970 6980
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT
AGAAGTTAGATGAGAGGTGAGGAGACCAGTTAGAAAGCCAATAAGCATTTCCAGATGAGAGATAATGGTTCT


6920 6930 6940 6950 6960 6970 6980
6990 7000 7010 7020 7030 7040 7050
TGAAATCCAATAGTGCCCAGGTCTAAATTGAGATGGGTGAATGAGGAAAATAAGGAAGAGAGAAGAGGCAAG
TGAAATCCAATAGTGCCCAGGTCTAAATTGAGATGGGTGAATGAGGAAAATAAGGAAGAGAGAAGAGGCAAG
6990 7000 7010 7020 7030 7040 7050
7060 7070 7080 7090 7100 7110 7120
ATGGTGCCTAGGTTTGTGATGCCTCTTTCCTGGGTCTCTTGTCTCCACAGGAGGAGCCATGGGGCACTACGT
ATGGTGCCTAGGTTTGTGATGCCTCTTTCCTGGGTCTCTTGTCTCCACAGGAGGAGCCATGGGGCACTACGT
7060 7070 7080 7090 7100 7110 7120
7130 7140 7150 7160 7170 7180 7190 7200
CTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA
CTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGA
7130 7140 7150 7160 7170 7180 7190 7200
7210 7220 7230 7240 7250 7260 7270
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACT
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACT
7210 7220 7230 7240 7250 7260 7270
7280 7290 7300 7310 7320 7330 7340
CCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGC
CCTTGATTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGC
7280 7290 7300 7310 7320 7330 7340
7350 7360 7370 7380 7390 7400 7410
CGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCC
CGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCC
7350 7360 7370 7380 7390 7400 7410
7420 7430 7440 7450 7460 7470 7480
TTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTT
TTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTT
7420 7430 7440 7450 7460 7470 7480
7490 7500 7510 7520 7530 7540 7550 7560
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCAT
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCAT
7490 7500 7510 7520 7530 7540 7550 7560
7570 7580 7590 7600 7610 7620 7630
AACCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTC
AACCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTC
7570 7580 7590 7600 7610 7620 7630
7640 7650 7660 7670 7680 7690 7700
TCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCA
TCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCA
7640 7650 7660 7670 7680 7690 7700
7710 7720 7730 7740 7750 7760 7770
CAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTG
CAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTG
7710 7720 7730 7740 7750 7760 7770
7780 7790 7800 7810 7820 7830 7840
GGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGTGTTTTTTC
GGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGTGTTTTTTC
7780 7790 7800 7810 7820 7830 7840



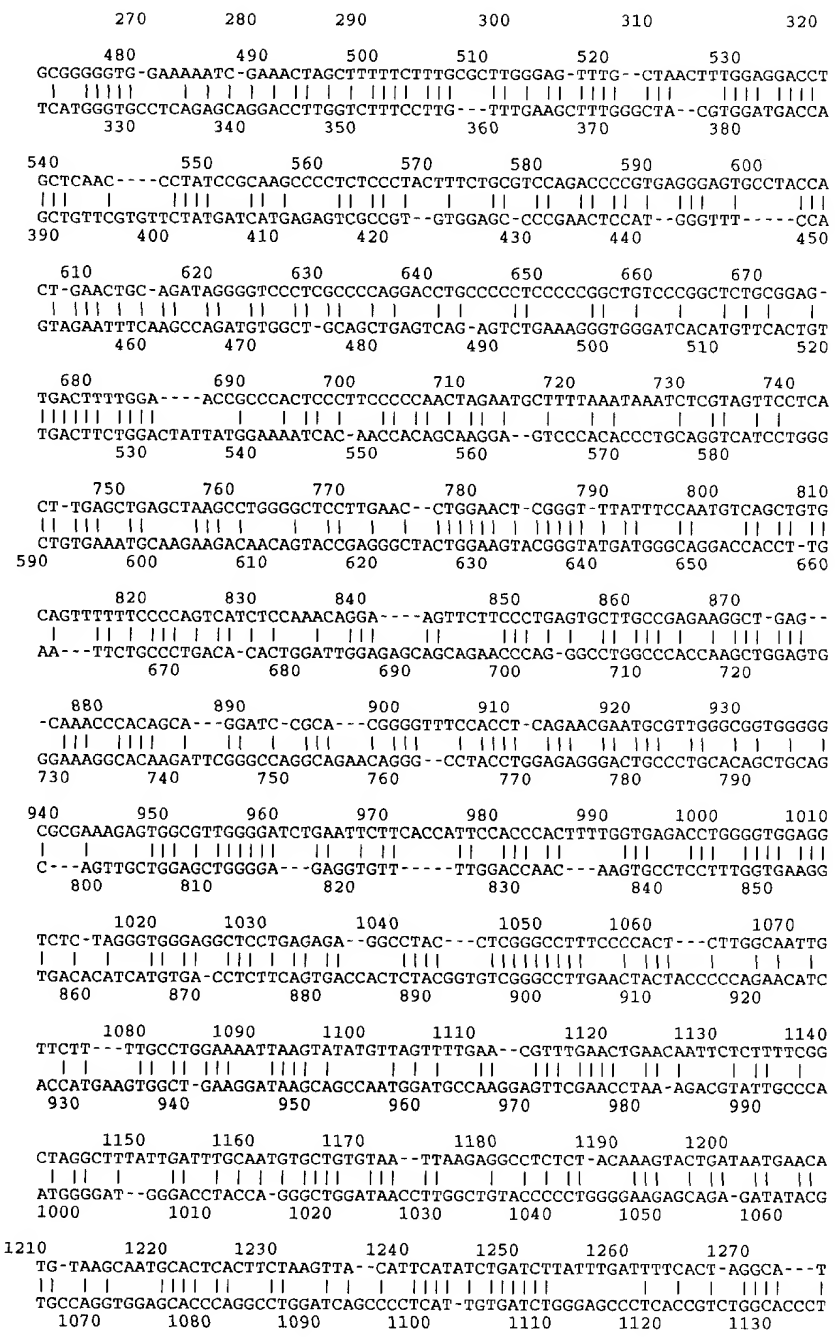


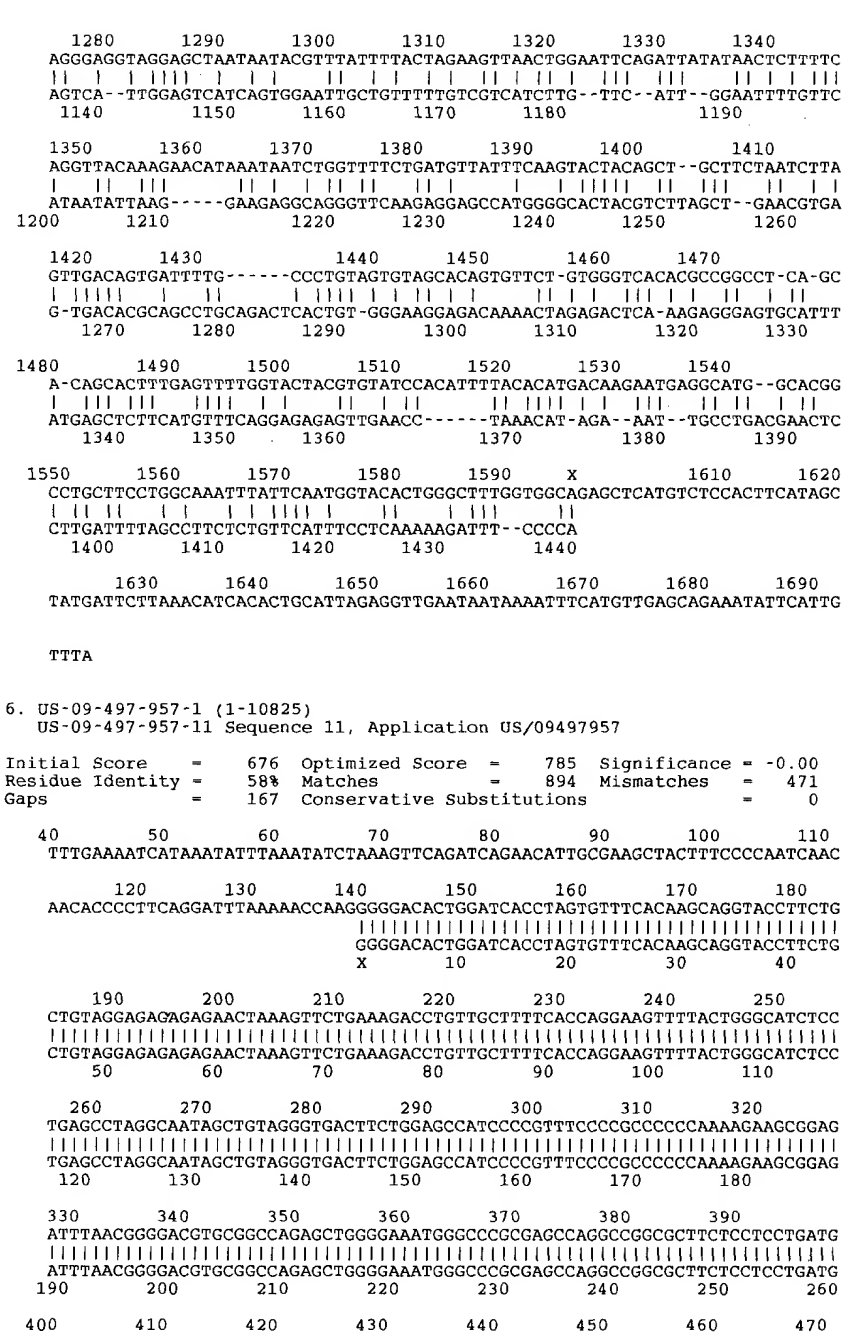




5. US-09-497-957-1 (1-10825)
US-09-497-957-9 Sequence 9, Application US/09497957

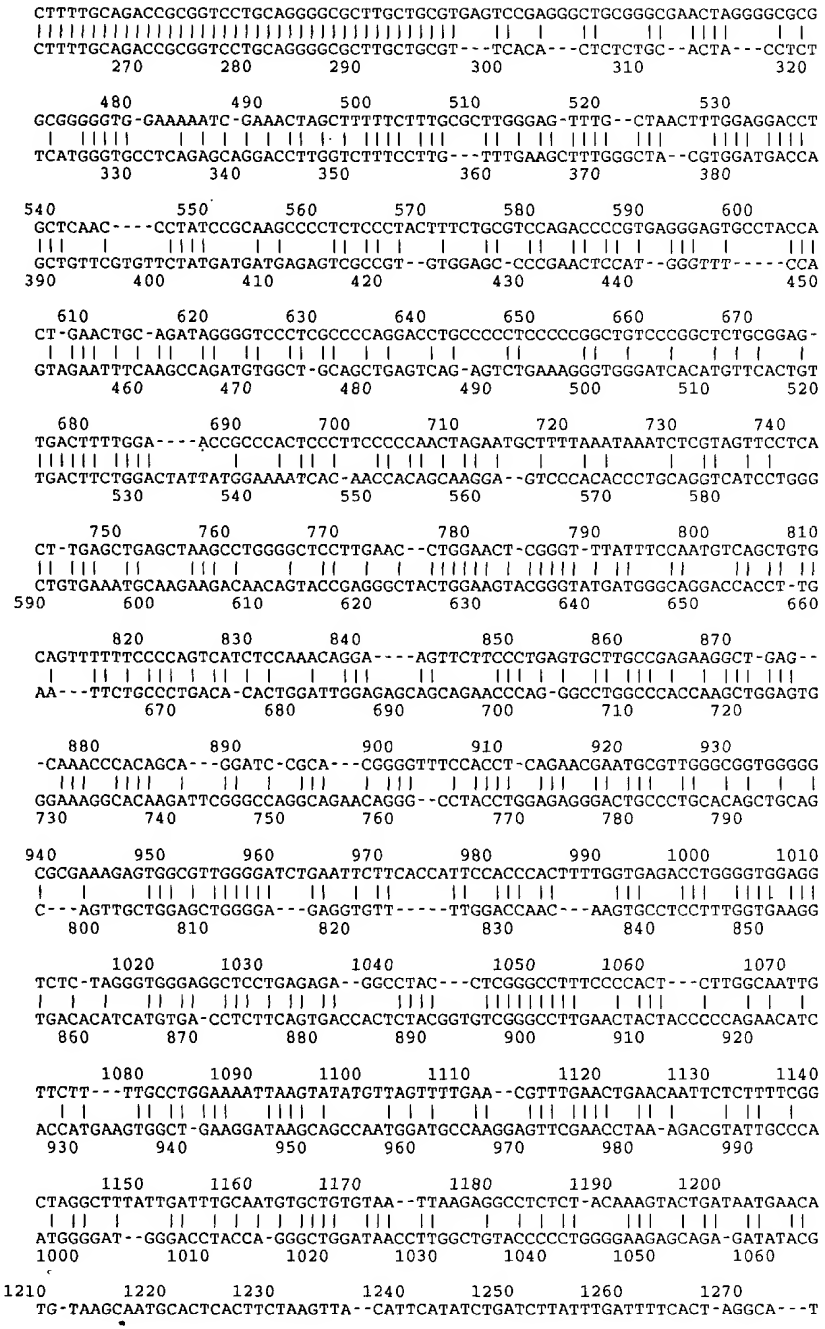
Initial Score	=	677	Optimized Score	=	786	Significance	= -0.00
Residue Identity	=	58%	Matches	=	895	Mismatches	= 470
Gaps	=	167	Conservative Substitutions				= 0

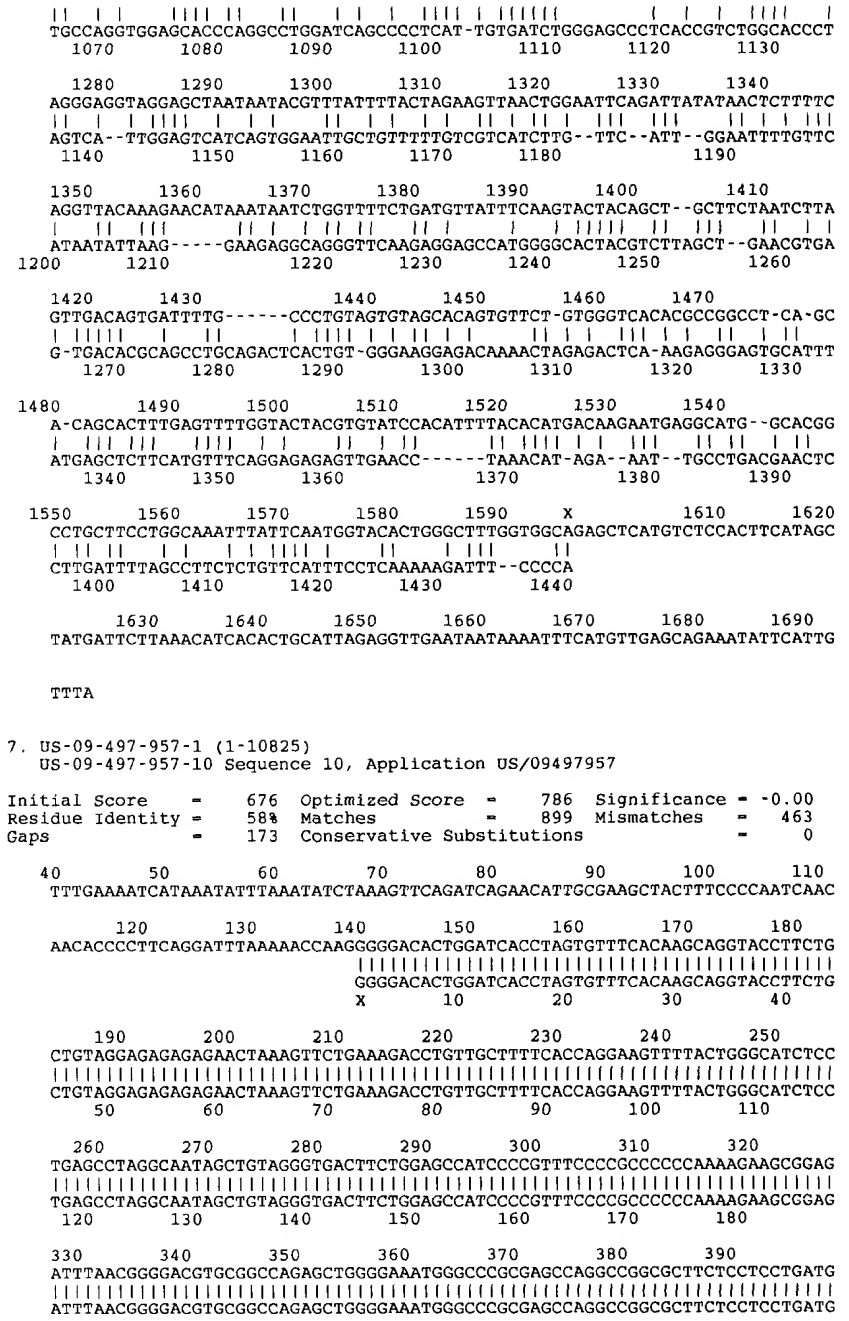




6. US-09-497-957-1 (1-10825)
US-09-497-957-11 Sequence 11, Application US/09497957

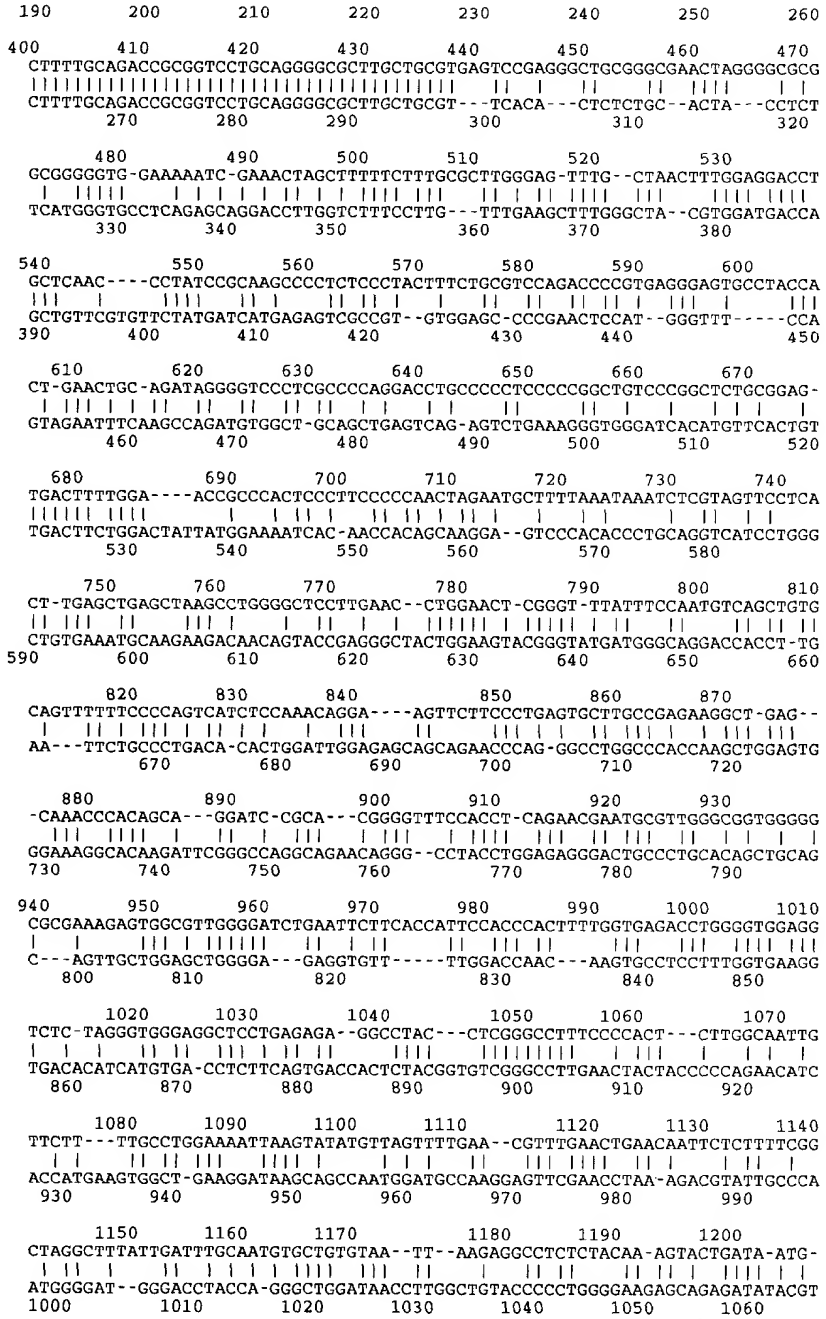
Initial Score	= 676	Optimized Score	= 785	Significance	= -0.00
Residue Identity	= 58%	Matches	= 894	Mismatches	= 471
Gaps	= 167	Conservative Substitutions		=	0

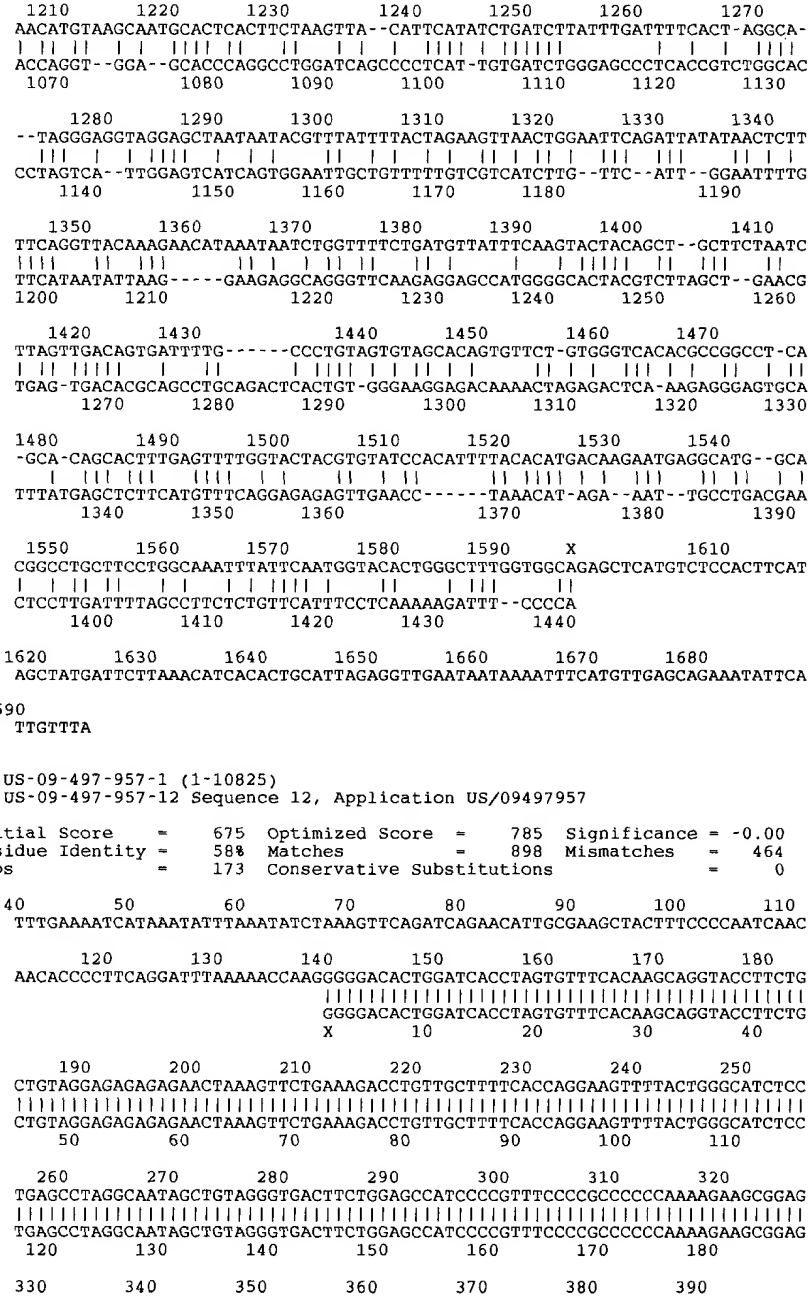




7. US-09-497-957-1 (1-10825)
US-09-497-957-10 Sequence 10, Application US/09497957

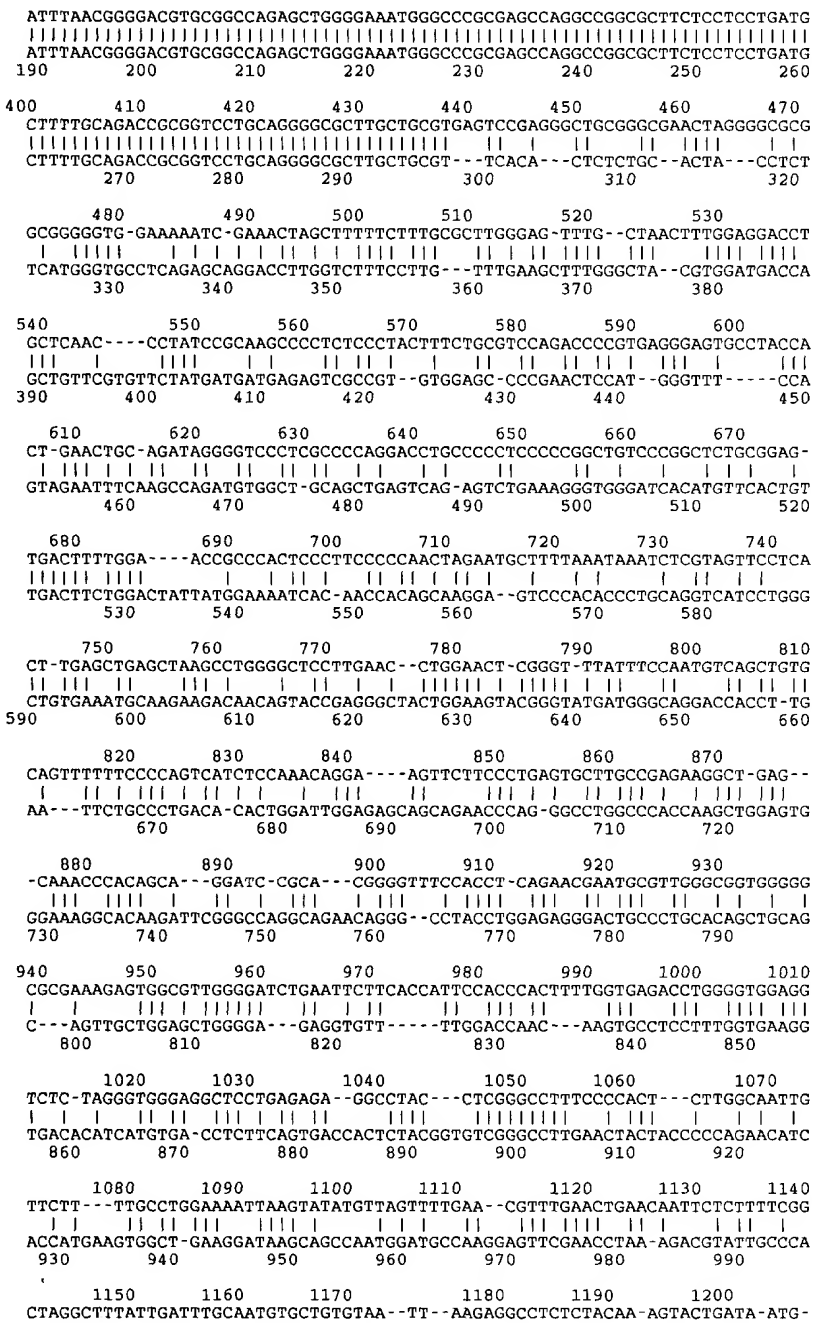
Initial Score	= 676	Optimized Score	= 786	Significance	= -0.00
Residue Identity	= 58%	Matches	= 899	Mismatches	= 463
Gaps	= 173	Conservative Substitutions			= 0

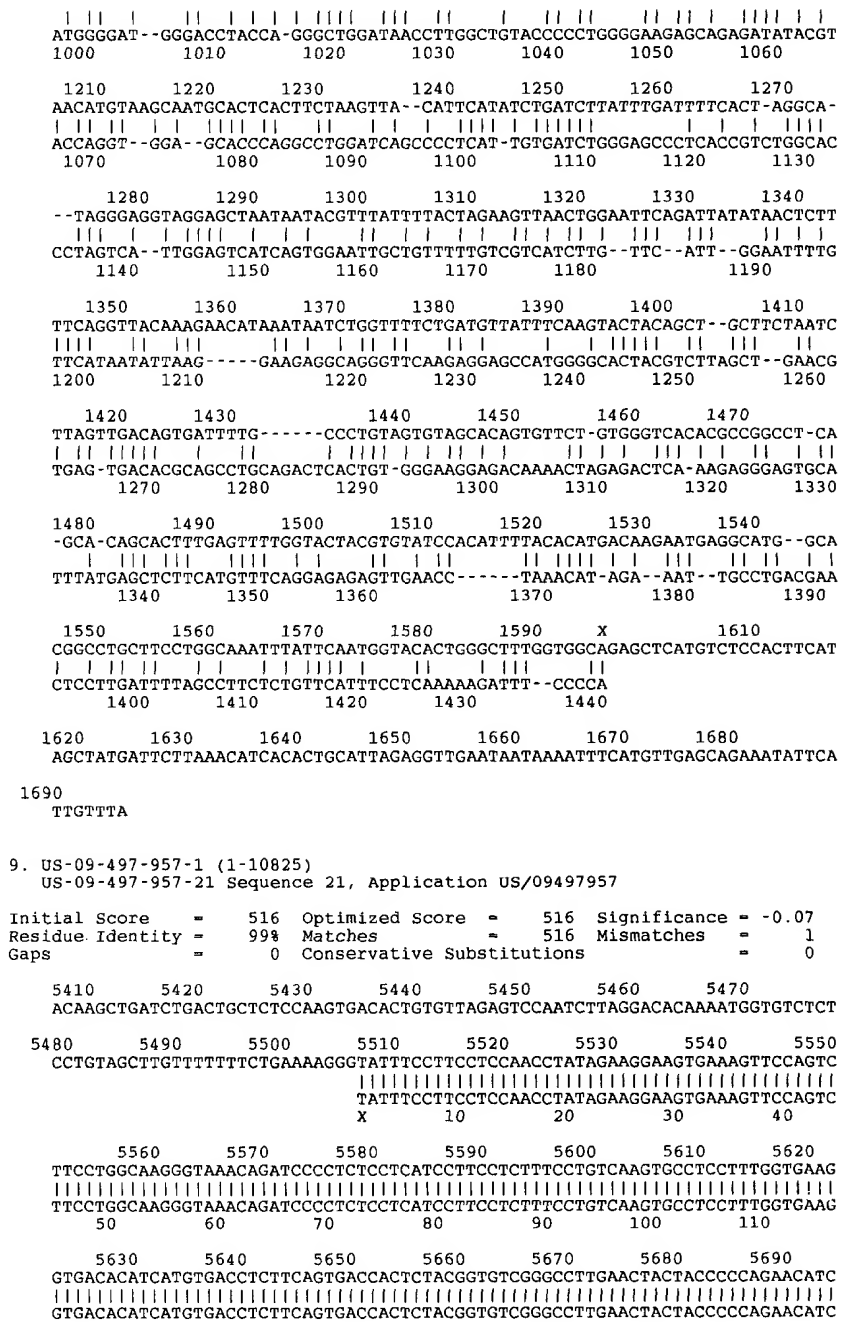




8. US-09-497-957-1 (1-10825)
US-09-497-957-12 Sequence 12, Application US/09497957

Initial Score	=	675	Optimized Score	=	785	Significance	= -0.00
Residue Identity	=	58%	Matches	=	898	Mismatches	= 464
Gaps	=	173	Conservative Substitutions				= 0





9. US-09-497-957-1 (1-10825)
US-09-497-957-21 Sequence 21, Application US/09497957

Initial Score	= 516	Optimized Score	= 516	Significance	= -0.07
Residue Identity	= 99%	Matches	= 516	Mismatches	= 1
Gaps	= 0	Conservative Substitutions			= 0

```
   120       130       140       150       160       170       180

     5700      5710      5720      5730      5740      5750      5760
   ACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAAT
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   ACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAAT
 190       200       210       220       230       240       250       260

   5770      5780      5790      5800      5810      5820      5830
   GGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTNCCAG
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
   GGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTACCAG
         270       280       290       300       310       320       330

5840       5850      5860      5870      5880      5890      5900      5910
   GTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGGAGCTGA
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   GTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGGAGCTGA
        340       350       360       370       380       390       400

         5920      5930      5940      5950      5960      5970      5980
   GAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGCTCTTTG
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   GAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGCTCTTTG
      410       420       430       440       450       460       470

       5990      6000      6010      6020  X   6030      6040      6050
   TTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCACCGTCTG
   |||||||||||||||||||||||||||||||||||||||||
   TTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTT
    480       490       500       510       X

     6060      6070      6080      6090      6100      6110      6120
   GCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGT
```

10. US-09-497-957-1 (1-10825)
    US-09-497-957-20 Sequence 20, Application US/09497957

Initial Score	=	516	Optimized Score	=	516	Significance	= -0.07
Residue Identity	=	99%	Matches	=	516	Mismatches	= 1
Gaps	=	0	Conservative Substitutions				= 0

```
   5410      5420      5430      5440      5450      5460      5470
   ACAAGCTGATCTGACTGCTCTCCAAGTGACACTGTGTTAGAGTCCAATCTTAGGACACAAAATGGTGTCTCT

 5480      5490      5500      5510      5520      5530      5540      5550
   CCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTTCCAGTC
                               |||||||||||||||||||||||||||||||||||||||||||
                               TATTTCCTTCCTCCAACCTATAGAAGGAAGTGAAAGTTCCAGTC
                               X       10        20        30        40

        5560      5570      5580      5590      5600      5610      5620
   TTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTGGTGAAG
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   TTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCCTGTCAAGTGCCTCCTTTGGTGAAG
      50        60        70        80        90       100       110

      5630      5640      5650      5660      5670      5680      5690
   GTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATC
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   GTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATC
    120       130       140       150       160       170       180

     5700      5710      5720      5730      5740      5750      5760
   ACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAAT
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   ACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAAT
 190       200       210       220       230       240       250       260

   5770      5780      5790      5800      5810      5820      5830
   GGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTNCCAG
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
```

```
   GGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTGCCAG
         270       280       290       300       310       320       330

 5840      5850      5860      5870      5880      5890      5900      5910
   GTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGGAGCTGA
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   GTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGGAGCTGA
        340       350       360       370       380       390       400

         5920      5930      5940      5950      5960      5970      5980
   GAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGCTCTTTG
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   GAAAATCTATTGGGGGTTGAGAGGAGTGCCTGAGGAGGTAATTATGGCAGTGAGATGAGGATCTGCTCTTTG
      410       420       430       440       450       460       470

       5990      6000      6010      6020  X   6030      6040      6050
   TTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTTGCTTTTTCTGTTTTAGAGCCCTCACCGTCTG
   |||||||||||||||||||||||||||||||||||||||||
   TTAGGGGGTGGGCTGAGGGTGGCAATCAAAGGCTTTAACTT
    480       490       500       510       X

     6060      6070      6080      6090      6100      6110      6120
   GCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGT
```

11. US-09-497-957-1 (1-10825)
    US-09-497-957-42 Sequence 42, Application US/09497957

Initial Score	=	39	Optimized Score	=	39	Significance	= -0.26
Residue Identity	=	97%	Matches	=	39	Mismatches	= 1
Gaps	=	0	Conservative Substitutions				= 0

```
        3760      3770      3780      3790      3800      3810      3820
   CTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTG

      3830      3840      3850 X    3860      3870      3880          X
   TTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATNATGAGAGTCGCCGTGTGGAGCCC
                                 |||||||||||||||||||| ||||||||||||||||||||
                                 AGCTGTTCGTGTTCTATGATGATGAGAGTCGCCGTGTGGA
                                 X       10        20        30        40

    3900      3910      3920      3930      3940      3950      3960
   CGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG

  3970      3980      3990
   GATCACATGTTCACTGTTGACTTC
```

12. US-09-497-957-1 (1-10825)
    US-09-497-957-41 Sequence 41, Application US/09497957

Initial Score	=	39	Optimized Score	=	39	Significance	= -0.26
Residue Identity	=	97%	Matches	=	39	Mismatches	= 1
Gaps	=	0	Conservative Substitutions				= 0

```
        3760      3770      3780      3790      3800      3810      3820
   CTGTCTCCAGGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTG

      3830      3840      3850 X    3860      3870      3880          X
   TTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATNATGAGAGTCGCCGTGTGGAGCCC
                                 |||||||||||||||||||| ||||||||||||||||||||
                                 AGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGA
                                 X       10        20        30        40

    3900      3910      3920      3930      3940      3950      3960
   CGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG

  3970      3980      3990
   GATCACATGTTCACTGTTGACTTC
```

13. US-09-497-957-1 (1-10825)
    US-09-497-957-30 Sequence 30, Application US/09497957

```
Initial Score    =     38  Optimized Score  =     38  Significance = -0.26
Residue Identity =    97%  Matches          =     38  Mismatches   =      1
Gaps             =      0  Conservative Substitutions               =      0

      5720      5730      5740      5750      5760      5770      5780
   ATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGG

     5790      5800      5810   X  5820      5830      5840      5850 X
   GCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTNCCAGGTGGAGCACCCAGGCCTGG
                                 |||||||||||||||||||| ||||||||||||||||||
                                 GGAAGAGCAGAGATATACGTACCAGGTGGAGCACCCAGG
                                 X       10        20        30        X

   5860      5870      5880      5890      5900      5910      5920
   ATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGGAGCTGAGAAAATCTATTGGGGGTTG

 5930      5940      5950
   AGAGGAGTGCCTGAGGAGGTAAT
```

14. US-09-497-957-1 (1-10825)
    US-09-497-957-29 Sequence 29, Application US/09497957

```
Initial Score    =     38  Optimized Score  =     38  Significance = -0.26
Residue Identity =    97%  Matches          =     38  Mismatches   =      1
Gaps             =      0  Conservative Substitutions               =      0

      5720      5730      5740      5750      5760      5770      5780
   ATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGG

     5790      5800      5810   X  5820      5830      5840      5850 X
   GCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATATACGTNCCAGGTGGAGCACCCAGGCCTGG
                                 |||||||||||||||||||| ||||||||||||||||||
                                 GGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGG
                                 X       10        20        30        X

   5860      5870      5880      5890      5900      5910      5920
   ATCAGCCCCTCATTGTGATCTGGGGTATGTGACTGATGAGAGCCAGGAGCTGAGAAAATCTATTGGGGGTTG

 5930      5940      5950
   AGAGGAGTGCCTGAGGAGGTAAT
```

15. US-09-497-957-1 (1-10825)
    US-09-497-957-43 Sequence 43, Application US/09497957

```
Initial Score    =     31  Optimized Score  =     31  Significance = -0.26
Residue Identity =    96%  Matches          =     31  Mismatches   =      1
Gaps             =      0  Conservative Substitutions               =      0

         3770      3780      3790      3800      3810      3820      3830
   GGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCT

      3840      3850      3860      3870      3880      3890 X    3900
   TTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATNATGAGAGTCGCCGTGTGGAGCCCCGAACTCCA
                                |||||||||| ||||||||||||||||||
                                TGTTCTATGATCATGAGAGTCGCCGTGTGGAG
                                X       10        20        30 X

     3910      3920      3930      3940      3950      3960      3970
   TGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATG

   3980      3990
   TTCACTGTTGACTTCT
```

16. US-09-497-957-1 (1-10825)
    US-09-497-957-44 Sequence 44, Application US/09497957

```
Initial Score    =     30  Optimized Score  =     30  Significance = -0.26
Residue Identity =    93%  Matches          =     30  Mismatches   =      2
Gaps             =      0  Conservative Substitutions               =      0
```

```
         3770      3780      3790      3800      3810      3820      3830
   GGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCT

      3840      3850      3860      3870      3880      3890 X    3900
   TTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATNATGAGAGTCGCCGTGTGGAGCCCCGAACTCCA
                                |||||||||| ||||||||||||||||||
                                TGTTCTATGATCATGAGAGTCGCCGTGTGGAG
                                X       10        20        30 X

     3910      3920      3930      3940      3950      3960      3970
   TGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATG

   3980      3990
   TTCACTGTTGACTTCT
```

17. US-09-497-957-1 (1-10825)
    US-09-497-957-31 Sequence 31, Application US/09497957

```
Initial Score    =     22  Optimized Score  =     22  Significance = -0.26
Residue Identity =   100%  Matches          =     22  Mismatches   =      0
Gaps             =      0  Conservative Substitutions               =      0

       220       230       240       250       260       270       280
   AGACCTGTTGCTTTTCACCAGGAAGTTTTACTGGGCATCTCCTGAGCCTAGGCAATAGCTGTAGGGTGACTT

     290       300       310   X   320       330     X  340       350
   CTGGAGCCATCCCCGTTTCCCCGCCCCCCAAAAGAAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGG
                                 |||||||||||||||||||||
                                 CAAAAGAAGCGGAGATTTAACG
                                 X       10        20 X

   360       370       380       390       400       410       420
   GAAATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGGCGC

 430
   TTGCTG
```

18. US-09-497-957-1 (1-10825)
    US-09-497-957-24 Sequence 24, Application US/09497957

```
Initial Score    =     20  Optimized Score  =     20  Significance = -0.26
Residue Identity =   100%  Matches          =     20  Mismatches   =      0
Gaps             =      0  Conservative Substitutions               =      0

       3640      3650      3660      3670      3680      3690      3700
   GACTGCAACTCACCCTTCACAAAATGAGGACCAGACACAGCTGATGGTATGAGTTGATGCAGGTGTGTGGAG

       3710      3720      3730      3740         X      3760      3770
   CCTCAACATCCTGCTCCCCTCCTACTACACATGGTTAAGGCCTGTTGCTCTGTCTCCAGGTTCACACTCTCT
                                |||||||||||||||||||
                                ACATGGTTAAGGCCTGTTGC
                                X       10        20

     3780      3790      3800      3810      3820      3830      3840
   GCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGA

   3850
   TGAC
```

19. US-09-497-957-1 (1-10825)
    US-09-497-957-18 Sequence 18, Application US/09497957

```
Initial Score    =     20  Optimized Score  =     20  Significance = -0.26
Residue Identity =   100%  Matches          =     20  Mismatches   =      0
Gaps             =      0  Conservative Substitutions               =      0

      3860      3870      3880      3890      3900      3910      3920
   GTTCGTGTTCTATGATNATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAG
```

```
  3930      3940      3950      3960      3970    X 3980      3990
   CCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTAT
                               ||||||||||||||||||
                              CTGAAAGGGTGGGATCACAT
                              X       10        20

4000      4010      4020      4030      4040      4050      4060      4070
   TATGGAAAATCACAACCACAGCAAGGGTATGTGGAGAGGGGGCCTCACCTTCCTGAGGTTGTCAGAGCTTTT


   CATC
```

20. US-09-497-957-1 (1-10825)
    US-09-497-957-13 Sequence 13, Application US/09497957

Initial Score	= 20	Optimized Score	= 20	Significance	= -0.26
Residue Identity	= 100%	Matches	= 20	Mismatches	= 0
Gaps	= 0	Conservative Substitutions		=	0

```
     5460      5470      5480      5490      5500      5510      5520
   CTTAGGACACAAAATGGTGTCTCTCCTGTAGCTTGTTTTTTTCTGAAAAGGGTATTTCCTTCCTCCAACCTA

   5530      5540      5550    X 5560      5570    X  5580      5590
   TAGAAGGAAGTGAAAGTTCCAGTCTTCCTGGCAAGGGTAAACAGATCCCCTCTCCTCATCCTTCCTCTTTCC
                               ||||||||||||||||||
                              TGGCAAGGGTAAACAGATCC
                              X       10        20

  5600      5610      5620      5630      5640      5650      5660      5670
   TGTCAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCC


   TTGA
```